

FIGURE 1

GGGGCTCGGCCAGCGCAGCGCTAGTCGGCTGGTAAGGATTACAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTGTGAAGTTGTAACAGAAAACTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCCTCAGTTCCCTCAGCCCTGTAATTGGACATCTGCTGCTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTTATGCATTG
CTACCATTTATGTCGTTATAAGCAAGTTCATGCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTGTACTTGAATACTGAGTTGTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTCAGACCATCCTTCCACAAATGCGAGCCAAACATGGCAAACAGTC
TTCTGGATCAGACTGTTGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTGCACAGTGGCAATTGGACTGATTAGAACAGAAACTCCATTGAAACCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAACGTTATGTCACTTCCCTTT
GGTTTTCTGACTTACATCGTGAATTTCAGAAAATTCTTACGGGTGGAAGCCAATTAC
TGGATTAACCCCTATGACACTGCACCTGCCCTATTACAATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAATATTCTGTAATGATTATGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGTAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAACACCTATGCTTACTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMQPKIHGKQVFIRLLLVIWCGVSAWSMLCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDfqKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGAGAGCCGAGTCCGGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGGCTGTGGCCCAGCGTGTGGCTCGGGAGTGGAGGGCAGGAGCCCTC
CTTACACTTCGCCATGATTCCTCCATGCACTCCAGCATCATGATTTACCCAGATACTATTTTTGG
GATTGGGTGGTTTCATGGCCATTGTTTAAAGACTATGAGATACGTCAGTATGTGTACAG
GTGATCTCTCCGTGACGTTGCATTTCTGCACCATTGTGACTCATCTGTAAATCTTAGG
AGTATTGAATAGCGCTCCGTTTTCACTGGAAATGAACCTGTGTTATTCTGTGATCCTGG
TTTCATGGGTGCCTTTTACTGGCTATTTGTGACAATATCCGACTACTGCCAAAACACGA
CTGCTTTTTCGTCTTTATGGCTGACTTTGTATTTCTGTGGAAACTAGGAGTCCCTTTCC
CATTCTCAGGCCAAACATGGATCTTACCATAGACAGCTCATCAGGCCGGTTGGGTGATTGGAG
TGACCTCAGGCTCTTTCTGGATTTGGGTGTCAACTGCCCCATACTTACTGATCTTACTC
CTCAGGGAATGTGACTGACGGATTCTAGCCCTGGGAACGGCGACGTGCCAAACCATGGATATGA
CATAAGCAAAAAAGAAAGGGATGGCAATGGCACGGGGAAACAATGTTCCAGAGGGGGAAGGTGCCATAACAA
AACCATCAGGTTCTGGGAATGAAAAAAGGTTTACCCACTTCAGCATCAGGAAGTGAAAATCTTA
CTTTATCAACAGGAAGGTGGATGTGTTTGGAAATTTCAAGGGCAGTCTTTCTGGAAACAGCTGATCAT
ATATGCTACCAAAGGGGAGAATAGAAATCTCCAAACCTTCAAGGGGAAATTTTATTCTGGTT
ACTTTTCTCTTTACTGTGTTTGGAAATTTCAGGCTACCAATTTGTTTTTGTCAGTGAGTTT
GGGGAAACGGATCTGTCAAGAGGGCATTGAGACTGTGAATATCTGGGAATCCAATTGTGAGTTT
GAAGTTTTGGGTCCAAACACATTTCTTCATCTGTGTGGGAATATCACGTCACATCCATCAGGGGATTT
TGCTGTCACTCTACCAGGTTCTGTGTTTATCCATCTGTCAGTACGTCTCCAAATGTCATTGTCCGT
CTTATTACCACAGAATGGGTCATTGTCTCTCCCTGTGTCTGTGATCCCGAATGAGTATGCCTTT
AGGAATACCGGCACCATACTCAGTGAAGTCCTGGGAACGTCACTTCATCACCGTTTGGTT
ATGTGATCTTCCGTGGTCAGCGCTCTCTCAGTCAACTCTCCCTCTATTTGGGTCAAAACAGGCACCACA
GAGAACCAATGGCACTTTGAACCTTAAGCCTACTACAGACGTTTGAGGCCAGGTGGTTCAAAATTT
GATATAAGGGGGGGAAATGGAACCAGGGCTGACATTTAAAACAAACAAATGTCATTGGGTACTT
ATTTTCACCTTCATAGCATACTCCCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACAATGAGGGGGAGAACTAACCAAGGACAATCTCAGCAGAGGACTCCGTGTGGATATGAGGGTGG
TGTAGGAGGGCGGAGAGGAGCCAAAGAAACTAAAAGGTGAAAAAACACATCAGTGAACCTGGGGCAAGACATGT
CTATGGGTAGGTGAGCCAAAACACGTGAGGATTTCCGTTTTAAGGTTCACATGGGAAAAGGTTTTAAGGTTT
CCTTGAGAATGTGATCATAAACATCAGAGACTGTGACAAAAAAAAAAAAAAAAAAAAGGGCGGGCGCG
ACCTCAGAGTGTCACCGTCAGAAGCTGGCCGGCAATGGCCCAATGTTTATTCGAGTCTTATATGA

FIGURE 4

MSFLIDSSIMITSQILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLIIVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGD
FPILSPKHGILSIEQLISRVGVIGVTMALLSGFGAVNCPTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKRMMARRTMFQKGEVHNKPSGFWMGIKSVTTSSAGSENLTLLIQQEVDALAEELSRO
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPPTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQMIMGY
FVSSVLLIRMSMPLERYRTIITEVLGELQFNFYHRWFVDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
AGAACCTCCATCCGGACTAGTTATTGAGCATTCGCTCTCATATCACCAAGTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCACT
TCCATCTGGACCACGAGGCTCTGGTCAAGGCTCTTGCGTGAGAAGAGCTTCCAGGT
GTCATGCAGAATTATGGGGATCACCCCTGTGAGCAAAAAGGCAGACCAGCAGCTGAATTTCACAG
AAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTGGCGGCAAGGACCAAGTGAACAGCC
TTGAAAGCTAGCTTGAAACTTGAGCTATGGCTGGGTGGGTCTGATTTGAGTTGGTGTCACTCTAG
GATTAGCCAAACCCCAAGTGTGGAAAAATGGGTGGGTCTGATTTGAGTTGGTGTCACTCCAGTGA
GCCGACAGTTGCAGCCTATGTTACAACACTCATCTGACTAAGTGGACTAACTCGTGCATTCAGAA
ATTATCACCACCAAAGATCCCATATTCAACACTCAAACACTGCAACACAAACAGAATTATTGT
CAGTGCAGTACACTCGGTGGCATCCCCCTACTCTAACAAACACTGCCCCCTACTACTACTCCTC
CTGCTCCAGCTTCACTTCTATTCCACGGAGAAAAAAATTGATTGTGTCAAGAGTTTATG
GAAACTAGCACCATGTCTACAGAAACTGAAACCATTGTTGAAATAAAGCAGCATTCAAGAATGA
AGCTGCTGGGTTGGAGGTCTCCCACGGCTCTGCTAGTGCTTGCCTCCTCTTGGTGTG
CAGCTGGTCTGGATTGGTATGTCAAAGGTATGTGAAGGCCTCCCTTTACAAACAAGAAT
CAGCAGAAGGAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCCTAA
TGAGGAATCAAAGAAAATGATAAAAACCCAGAAGAGTCCAAAGAGTCCAAAGCAGCAGCAG
GATGCCCTGAAACTGAAGTT**TAGATGAGACAGAAATGAGGAGACACACTGAGGCTGGTTCTT**
CATGCTCCTTACCCCTGCCCCAGCTGGGAAATCAAAGGGCCAAGAACCAAGAACAGTCCA
CCCTGGTTCTTAACGGTCACTGGAAATCAGCTCAGGACTGCATTGGACTATGGAGTGCACCAAGAAC
GCCCTCTCCTATTGTAACCTGTGGATCTATCCTCCTACCTCCAAAGCTTCCACGGCCT
TTCTAGCCTGGTATGCTTAATAATATCCCACGGAGAAAGGAGTTTGCAAAGTGCAGGAC
CTAAACATCTCATCAGTATCCAGTGGAAAAAGGCCCTCTGGCTGTGAGGCTAGGGTTG
AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTACTGATTCGCGCAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAAACACGTATCCCACCTGACATGCTCTGAGCCCCGGTAAGAGCAGAAC
GGCAGAAAAGTTAGCCCTGAAAGCCATGGAGATCTCATAACTTGAGACATACTCTGTAAA
GCTAAAATAAAAGAAATAGAACAGGCTGAGGATACAGTACACTGTCAGCAGGGACTGAAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGGAAATCACTGTTAGAACACACACA
CTTACTTTCTGGTCTTACCACTGCTGATATTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAAATCTTATAAATTCTATTCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGTGCAAGGTATTACACTCTGTAATTGAATTATTCTCAAAAATTGCACATAGTAG
AACGCTATCTGGAGCTATTCTGAGTTGATATTCTAGCTTACTCTCAAAACTAAT
TTTATTCTGAGACTAATCTTATCTATTCTAATATGGCAACCAATTATAACCTTAATT
TATTATAACATACCTAACAGAAGTACATTGTTACCTCTATATACCAAAAGCACATTAAAAGTGC
ATTAAACAAATGTACTAGCCCTCCTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATT
TGTGACAAAAAATTAAAGCATTAGAAAACCTT

FIGURE 6

MARCFSLVLLTSIWTRLLVQSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAASFNEAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCCGCGCTCCCGCACCGCGGCCGCCACCGCGCCGCTCCCGCATCTGCACCCGAGCCCGC
GGCCTCCCGCGGGAGCGAGCAGATCCAGTCAGCGCCAGCGCAACTCGGTCCAGTCGGGGCG
CGGCTCGGGCGCAGAGCGGAGATTCCAGCGCTTGGGCCACCTGCTGTGCCTGCTGCGG
CGGGCGTCCCCACGGGCCCCCGCCACGGGACCTCGGCTCAGTCAGAAGCCGGCCG
GCTCTCAGCTACCGCAGGGAGGACCCCTAATGAGATGTCAGGAGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGGGTGGAGAGATGGAGGAGAAGAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACATTACCTCCAGCTATCACAATGAGACCAACAGAC
ACGAAGGTTGGAATAATACCATTCATGTCACCGAGAAAATTACAAGATAACCAACAACCAGAC
TGGACAAATGGCTTTCAAGAGACAGTTACATCTGTCAGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTAACGCCAGTTGCCAGTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGCCACCAAGGGGAGCAATGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCGGGCTGTGTCAGGAGGCCTGCTGTTCCAGAGAGGCCTGCTGTC
ACACCCCTGCCGTGGAGGGCAGCTTGCATGACCCGCCAGCCGGCTCTGGACCTCATCAC
CTGGGAGCTAGACCTGATGGAGCCTGGACCGATGCCCTTGCCAGTGGCTCCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGCAAGCCGACCTCGTGGGAGCCGTGACCAAGATGG
GAGATCCTGCTGCCAGAGAGGTCCCCATGAGTATGAAGTGTGCAGCTTCATGGAGGAGGTGCG
CCAGGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCAGTGGGGAGCCTGCC
CCGGCGCTGCACTGCTGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGTAGATGTGCAA
TAGAAAATAGCTAATTATTCAGGTGTGCTTAGCGTGGCTGACCAGGTTCTCCTA
CATCTTCTCCAGTAAGTTCCCTCTGGCTTGACAGCATGGTGTGCAATTGTTCAAGCT
CCCCCAGGCTGTCAGCTCACAGTCTGGCTGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCAGATTGGCTGCTTGCCTCTACCAGTTGGCAGACAGCCG
TTTCTTCTACATGGCTTGATAATTGTTGAGGGGAGGAGATGAAACATGTGGAGTCTCC
TGATTGGTTGGGAAATGTGGAGAAGAGTGGCCCTGTTGCAACATCAACCTGGCAAAATG
CAACAAATGAATTTCAGCAGTTCTCCATGGCATAGGTAAGCTGTGCTTCACTGTTGC
AGATGAAATGTTCTGTTCCAGCTCACAGTCTGGCTGGAGAGTCAGGCAGGGTTAAACTGCA
TACCTCTGCCCCAGGGCAGCATTTCTACATCCAAGATCAATTCCCTCTCAGCACAGCTGGG
AGGGGGTCAATTGTTCTCCTCGCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCT
CAAGTCACACAGCTAGTGAAGACCAAGCAGCAGTTCTACAGGGTGTGACTCTAAGCTCAG
CTCCACTACCCACACCAGGCTTGGGCCACAAAGTGCCTCCAAAAGGAAGGAGAATGGGAT
TTTCTTGGGCATGCACATCTGGAATTAAAGTCAAACTAATTCTCACATCCCTCTAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGAGCCGCTCTCATGAAGACATGAT
ATTGACACTGTCCTCTGGCAGTTGCATTAGTAACCTGGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAACAGTACTTAGGTAATTGTTAGGGCAGGATTATAATGAAATTG
AAAATCACTTAGCAGCAACTGAAGACAAATTCAACCAACAGTGGAGAAAATCAAACCGAGCAGGG
TGTGTGAAACATGGTTGTAATTGCACTGCAACTGCAACACTGAACCTACGCCACTCCACAAATGAT
TTTCAAGGTGTCTGGACTGTTGCCACCATGTTACATCCAGAGTTCTAAAGTTAAAGTGTCA
CATGATTGTATAAGCATGCTTCTTGAGTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAATCACTCAACTGCAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAAKASSEVNLANI.PPSYHNETNTDTKVGNNNTIHVHREIHKITNNQTGQMVSE
TVITSVGDEEGRSHECIIDEDCGPSMYCQFASFQYTCQPCRQMLCTRSECCGDQLCVWGHC
TKMATRGSGNTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACGCGTGGGGCTGTGAGAAAAGTCCAATAATACATCATGCAACCCAC
GGCCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGTACTCATCCAAG
GCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGTCTGGGGCTCTCTGGACCCCTT
AACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTGCCTCCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGTCATTGGCATTGGAGCCCTCATCCTGACCCTTGTGCAGATAGCCGGTCATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGCCCGTGCATCATGTGCTGTT
CAAGTGCCTGCTCTGGTGTCTGGAAAAAATTATCAAGTCCATAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTGTCAGCCAAAATGCCTCATGCTACTCATGCGAAC
ATTGTCAAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTTGGAAAGCTGCTGGT
GGTCGGAGGCGTGGGGTCTGTCTTCTTCTCCGGTGCATCCGGGGCTGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCCTAT
GTCATGCCAGGGCTTCTCAGCGTTTGGCATGTGTGGACACGGCTCTCCCTGCTTCT
GGAAGACCTGGAGCGAACACGGCTCCCTGGACCGGCCACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACGCTCGG
CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCATCCAACCTCACTCGCTTACAGGT
CTCCATTGTGGAAAAAAGGTTAGGCCAGGCGCCGTGGCTCACGCCGTAAATCCAACACT
TTGAGAGGCTGAGGCCGGGATCACCTGAGTCAGGAGTTCGAGACCGCCTGGCCAATGGTGA
AAACCTCCGTCTCTTAAACAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCA
GCTACTCGGGAGGCTGAGGCAGGAGAACGCTGAACCCGGAGGCAGAGGTTGCAGTGA
GATCGGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAA
AAGATTTATTAAAGATATTGTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVVLDKVTDLFFFKGKLLVVGGVGVLSSFFFFSGRIPGLGKDFKSPHLNYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GGCCCGCGCCCGCGCGCCGGCGCCGAAGCCGGGAGCCACCGCCATGGGGCCTGCCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCTGCACTCTGTGAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCTCATCTTCACGTTCTCCTCTCCGGGTGCTGGTGTCCA
TCATTATGCTGAGCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTTGAGGAGGGGCC
GGGATCCCCACCGCCTGCAGGGCACATCGACTGTGGCTCCCTGCTGGCTACCGCGCTGCTACCG
CATGTGCTCGCACGGCGCCCTCTTCTTCTTCAACCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGTTTGGTTCTTAAGTTCCTGATCCTGGTGGCCTC
ACCGTGGGTGCCCTCTACATCCCTGACGGCTCCCTCACCAACATCTGGTTCTACTTCGGCGTGG
CTCCTCCTCTTCATCCTCATCCAGCTGGTGTGTCATCGACTTGCACACTCTGGAAACAGCGGT
GGCTGGCAAGGGCAGGGAGTGCAGTCCCGTGCCTGGTACGCAGGCCTCTCTTCACTCTCCTC
TTCTACTTGCTGTCATCGCGGCCGTGGCGTGTGGTACAGTACTACACTGAGGCCAGCGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTAACCTCACCTCTGTGTCTGCGTGTCCATCGCTGCTGCTGC
CCAAGGTCCAGGACGCCAGCCAACCTGGGCTGCTGAGGCCCTGGTCATCACCCCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGCAACGAGACAGTTGTGGCAGGCCAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
TGGCCTCATCATCTCCTCTGTGCACCCCTTCATCAGTCTGCGCTCCTCAGACCACCGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTAGGGCCGGGCTTGACAACAGCAGCAGGCCGTACCTACAGCTACTCCTTCTCC
ACTTCTGCCTGGTGCTGGCCTCACTGCACGTATGATGACGCTCACCAACTGGTACAAGCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCGAACCGCGACTTCAGCTGAGGCTCA
CAGCCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGCTCGGTGACAGCCAACCTGCCCG
CCCACACCAATGCCAGGCTGAGCCCCCAGCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGG
CTTCTAGTCGTAGTGCCTCAGGTCCAGGGTCCAGGGAGCATCAGGCTCTGCGAGGCCACCTGCC
ACCCACACGGTGGAGCTGCCCTTCCCTCCCTCCCTGGTGCAGAGGCCACCTGCC
AGGGCTCCCTGTGCCTCAGGCTCCAGGGAGCGGGGCTGCTGGAGAGAGCGGGAACTCCCAC
TGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGACCCCTGCC
GACTTCGTGCCTACTGAGTCTAAGACTTTCTAATAAACAGCCAGTGCAGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGIFTNIWFYFGVVGSFLFILQLVLLIDFAHAWNQRWLKGAE
ECDSRAWYAGLFFFLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIIFLLCTLFISLRRSDHRQVNLSMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICASWAGLLYLWTLVAPLLLNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCAGGAACCACCGTTAAGGTGCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGGCCAGAGCTCCATGCTCTCTGCGCAATATCCATTCCATCAACCCCACACAA
CTCATGGCCAGGATTGAGTCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTCTGTTGTTGTCACCTTGACCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGGAGGTGATGCACTGACTACTATTCTCATAT
TTGATATATTCCTGGCAGTTTCGATTTAAAGTGTAAATACCTGCATATGCTGTGTCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAG
TGATCCTTCGAAGCTTTCTCAAGGGCTTTGGCTATGTGCTGCCATCATTTCATTTCATC
CTTGCTGGATTGAGACGTGGTCTGGATTTCAAAGTGTACCTCAAGAACGAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTCAGAGAGGGCAGCACTTACCTGGTGTCTTGATG
GTCAGTTTATTCCCCCTCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAACAGGACAGT
GAGAAACCACTTTAGAACTATGAGTACTACTTTGTTAAATGTGAAAACCCACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTACGTC
CACTGCTGGTTATTGAACAGCTAATAAGATTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTGATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTACAAATAGGTGAAGAAAGTCTGTGCTGATTCTTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTAGTAAGCAAGATACTTTTATTCAATTCAAC
AGAATGGAATTTTTGTTCATGTCAGATTATTTGTATTCTTTAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTGCTTTGACAGTTAAAAAGTGTAAATAAAATCTG
ACATGTCAATGTGGCTAGTTTATTCTGTTGCTGATTGTGTTGATTGTCAGTGGCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAGGACATGGTACCTGAAATGTTCAAAACTCTT
TTATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCAAAACTCTTATTGAAATGC
ACAAAATGACTAAACCATTCAATCATGTTCTTGCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIIHSINPTQLMARIESTYEGREKKISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIPLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPISFILAWIETWFLDFKVLPOEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPSEAGSSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCCTCGGGCCCCGACCCGCCAGGAAGACTGAGG
 CCGCGGCCCTGCCCGCCGCCGCTCCCTGCCGCCGCCGCCCTCCCGGGACAGAAGATGTGCTCCAG
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGTGCAGGGCTGCCCAT
 CCGGCTGCCAGTGCAGGCCACAGACAGTCTTCTGCACTGCCGCCAGGGGACCACGGTGCC
 CGAGACGTGCCACCCGACACGGTGGGGCTGACGCTTCTTGGAGAACGGCATCACCATGCTCGACGC
 AGGCAGCTTGCCTGGGGCTGCCGCCGCTGACGCTCCTGGACCTGTCACAGAACAGATGCCAGCC
 TGCCAGCAGGGTCTTCCAGCCACTGCCAACCTCAGCAACCTGGACCTGACGCCAACAGGCTG
 CATGAAATCACCAGAACCTTCCGTGGCTGCCGCTCGAGCGCTCTACCTGGCAAGAA
 CCGCATCCGCCACATCCAGCCTGGTGCCTCGACACGCTCGACCGCTCTGGAGCTCAAGCTGC
 AGGACAACGAGCTGCCGGCACTGCCCGCTGCCCTGCCGCCCTGCTGCTGGACCTCAGC
 CACAACAGCCTCTGGCCCTGGAGCCCGCATCCTGGACACTGCCAACGTGGAGGGCGCTGCC
 GGCTGGTCTGGGCTGCCAGCAGCTGGAGGGCTCTCAGCCGCTGCCAACCTCACGACC
 TGGATGTGTCGACAACCAGCTGGAGCGAGTGCACCTGTGATCCGAGGCCCTCCGGGGCTGACG
 CGCCTGCCGCTGGCCGGAACACCGCATTGCCAGCTGCCGCCAGGGACCTGGCCGGCTGG
 TGCCCTGCCAGGAGCTGGATGTGACCAACCTAACGCCCTGAGGCCCTGCCCTGGGACCTCTGGGG
 TCTTCCCCCGCCTGCCGCTGCCAGCTGCCAACCTTCAACTGCCGTGCCCCCTGAGC
 TGGTTTGGCCCTGGGTGCCGAGAGCCACGTACACTGCCAGGCCCTGAGGAGACGCC
 CTTCCCGCCAAAGAACGCTGCCGGCTGCTCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG
 CCACCAACACAGCCACAGTGCCACCCACAGGCCCCGTTGGTGCAGGGAGGCCAACGCC
 TCTAGCTTGGCTCCTACCTGGCTAGCCCCACAGGCCGCCACTGAGGCCAGGCC
 CACTGCCCAACCGACTGTAGGGCCTGCCCCCAGCCCAGGACTGCCACCGTCCACCTGCC
 ATGGGGCACATGCCACCTGGGACACGGCACCACCTGGCGTGTGCCCCGAAGGCTTCACG
 GGCGTACTGTGAGAGCCAGATGGGCAAGGGGACAGGCCAGGCCCTACACCAGTCACGCC
 GCCACACGGCTGCCGCTGACCCCTGGGATGCCAGGGGGCTGCCGCTGGGGCTG
 AGCGCTACCTCCAGGGGAGCTCCGTGACGCTCAGGAGCCTCCGTCTCACCTATCG
 GGCCCTGATAAGCGGCTGGTGACGCTGCCACTGCCCTGGCTGAGTACACGGTACCC
 GCTGCCGCCAACGCCACTACTCCGCTGTGTCATGCCCTTGGGGCCGGGGTGCCGG
 GCGAGGAGGCCCTGCCGGGAGGCCATACACCCCCAGCCGTCACCTCAACCAACAGGCC
 CAGGCCGCCGAGGGCAACCTGCCGCTCCTCATGCCGCCGCCCTGCCGCGTGTCTGCC
 GCTGGCTGCCGGGGCAGCCTACTGTGTGCCGGGGGGGCCATGCCAGCAGGCC
 ACAAAAGGGCAGGTGGGGCAGGGGCTGGGGCTGGGAACCTGGAGGGGAGTGAAGG
 CCAGGGGAGGCCAACAGAGGGGGCTGGAGAGGCCCTGCCAGGCCGCTGAGTGT
 ACTCATGGCTTCCCAGGGCTGCCCTCAGTCACCCCTCCAGCCTGAGTACATC
 GAGAGAGACAGGGCAGCTGGGGCCGGGTCTCAGCCAGTGCAGATGGCC
 ACACCACTGAAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCA
 GAGACAGGGCTGTGACCTGAGGAGATGCTGAGTGTGAC
 GAGCCCTGTTCCCTCTGGACCTCGGTCTCCCTCATCTGTGAGATGCTGTG
 GAGGACAGTGCAGTC
 CTAACGTCCCCAGAACCGAGTGCCTATGAGGACAGTGTGCCCTGCCCTCC
 GCAGTC
 CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCC
 CTGGGTCTGCTGGCTCTCC
 TCCAGGCGGACCCCTGGGGCCAGTGAAGGAAGGCTCC
 GGAAAGAGCAGAGGCCAGAGGGAGAGGCC
 GGCTGTGACTCTAGTCTGGCCCCAGGAAGCGAAGGAACAAAAG
 ACTGGAAAGGAAGATGC
 TTTAGGAACATGTTGCTTTAAATATATATTATAAGAGATC
 CTTCCATTATTCT
 GGGAAAGATGTTTCAAACCTCAGAGAACAGGACTTGGTTTG
 TAAGACAAACGATGATGAA
 GGCTTTGTAAGAAAAAATAAAAGATGAAGTGTGAA

FIGURE 16

MCSRVPLLLPLLLLALGPGVQGCPGQCQSPQTVFCTARQGTTVPRDVPPDTVGLYFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRPLRLLLDLSHNSLIALEPGIILDANVE
ALRLAGLGLQQLDeglFSRLRNLDLVDSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLALPGDLSGLFPRRLRLLAARNPFCNCVCPFWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGGTRPSPTP
VTPRPPRSLTGLIEPVSPSTSRLVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPGPGRVPEGEACGEAHTPPAVHSNIAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAVYCVRRGRAMAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGPGLQSPHLAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCAGGCCGGTGGCTGAGTCGTGGCAGAGGCGAACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGCGGTGCTGAGCTGGCCTGGCGTCCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCAAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGATTCAAAGAATCTGA
ATTAGAACCTCTATTCAAAGAAGAGAACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AACATATCAGCTTCTAGAGTCTCAAACAGAAAACAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACAGCTTGACGCCATTGAAGGCACAGCACATGGGAGCCCTGCCACTTCCCTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAAACCTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACGTGAATGAAAATCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAACGATATCGGTATCTCAAAGGCAGCAAGCATGAAACATA
CCAAAGCCCTGGAGAGAGTGTATATGCTTTATTGGTATTACTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTGAGAAGCAGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTGG
CTTCTGTATGCCCTGGACTTGGTGTAAATTCAAAGTCAGGCAAAGGCTCTGTATATTACAT
TTGGAGCTCTGGGGCAATCTAATAGCCCACATGGTTTGGTAAAGTAGACTTTAGTGGAAAGGCT
AATAATATTAACATCAGAAGAATTGTTGTTATAGCGGCCACAACCTTTCAAGTCTGTTATGGATAT
CAGATTGCTGTATTAAGACCAATATTCAAGTGAACCTCCTCAAATTCTGTTAATGGATAT
AACACATGGAATCTACATGTAATGAAAGTGGTGGAGTCCACAATTTCCTTAAATGATTAG
TTGGCTGATTGCCCTAAAAGAGAGATCTGATAATGGCTCTTTAAATTCTGAGTTG
GAATTGTCAGAATCATTTTACATTAGATTATCATATTAAATTCTTCTTGTGTTCA
AAATTGTAATGGTGGCTATAGAAAAACACATGAAATATTACAAATTGGCAACATGC
CCTAAGAATTGTTAAAATTCTGGAGTTTGTGCAGAATGACTCCAGAGAGCTACTTCTG
TTTTTACTTCATGATTGGCTGTCTCCCATTTATTCTGGTCAATTGCTAGTGCACACTGT
GCCTGCTTCCAGTAGTCTCATTCCATTTGCTAATTGTTACTTTCTTGTCAATTGG
AAGATTAACCTTTAAATAAAATTATGCTAAGATTAAAAAA
AAA

FIGURE 18

Important features:
Signal peptide:
amino acids 1-21

N-glycosylation sites.
amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.
amino acids 220-228

N-myristylation sites.
amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.
amino acids 267-271

Microbodies C-terminal targeting signal.
amino acids 299-303

Type II fibronectin collagen-binding domain protein.
amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.
amino acids 101-119

FIGURE 19

AATTCA GATTTAAGCCCATTCTGCAGTGGATTTCATGAAC TAGCAAGAGGACACCATTTCTT
GTATTATA CAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAATGCTTTTGGGTGCTAGG
CCTCTTAATCCTCTGGTTTCTGTGGACTCGTAAAGGAAA ACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTGGAAACTTGGCAGCCAGAACTTTGAT
AAAAAGGGATTTCATGTAATCGCTGCCTGTACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTCTGGATGTGACCAGCCAGAGAAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGTT
CCCGCGTGCTGGCTCCCCTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTGCCTGCAATGTTGGAGGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGACATGAAAGCTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAACAAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAAA
AGTCTAGACAAACTGAAAGCAATAATCCTATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTAAACAAGTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTGCAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCAATGTCTCCAGGCTATGA
AATTGGCCGATTCAAGAACACATCTCCTTTCAACCCCATCCTTATCTGCTCCAACCTGGACT
CATTAGATCGCTTATTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCAGGGT
CCCTGCTCAAGTTTCTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTTAGGCTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCATTCAAATGA
TCTTACCGTGGCCTGCCCATGCTTGGTCCAGCATTTACAGTAACCTGTGAATGTTAAGT
ATCATCTTATCTAAATATTAAAGATAAGTCAACCCAAAAAAA
AAAAAAAAAAAAAA

FIGURE 20

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGC GGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTGGGCTTGTGCTCGGCG
CACTCGTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTCTTCTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATCCAAATGGATGATGTTGAAGTTGTTATACAATTGA
CATTAGAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGACAAATTCCGT
CGTCATTCA GATCAGATCATGACGTTAGAGAGGGCTGCTCACAAAAC TTGCAAGGAGCATT
TTCAAACCAAGACCTGTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTACTC
ATCGACTGGAACATTCTTATATAAACCTCAAAAGGACTTTACAGGGTACCTTAGTGGT
GCCAATCTGGGATGTCTGAACA ACTGGGTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTGAGAAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATT CAGG CAGCAAGAGAGAAGAACATCCAAAAGACCCCTCAGGAGA
ACATTTTCTTGTCA GG CATTACGGACCTTTCCAATTCTGAATTCTCATT CATGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTAACTACAACCACCATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAAGCCTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTGT
GATACACAAGACAAACGATCTAAAGCAAATCTGGTAGTAGTAAACCAAGATAAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTGGTGAATATTCA CGGTCTC
CTACATTTGATCTTAACCTTACAAGGAGATT TTTTATTGGCTGATGGTAAAGCCAAAC
ATTCTATTGTTTACTATGTTGAGCTACTTGCAGTAAGTCATTGTTTACTATGTTCA
TGTTGCAGTAATACACAGATAACTCTAGTGCATTACTTCACAAAGTACTTTCAACATCA
GATGCTTTATTCCAACCTTTTCACTTCACTAAGTGTGAGGGAAAGGCTACACAG
ACACATTCTTACAAGTGAAGGAGGATTGAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
TAGGAAAGACAAGTCAGGAGGATTGAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCATGTCATTAAAAAATGGAAAAGCAAGAATAGCCTATTTCAAAATATGGAAA
GAAATTATGAAATTCTGAGTCATTAAATTCTCCTTAAGTGTAACTTTTAAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAA

FIGURE 22

Important features:
Signal peptide:
amino acids 1-19

N-glycosylation sites.
amino acids 75-79, 322-326

N-myristoylation site.
amino acids 184-154

Growth factor and cytokines receptors family.
amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGCGGGAGGGCAGAGTCAGCCGAGGCCAGTCCAGCCGGACAGCGCAGGGCAGCCAA
GCAGCGCGCAGCGAACGCCGCCGCCACACCCCTCGCGTCCCCGCGCGCTGCCACCCCTCCCTCCCT
GCGTCCCCGCCCTGCCGGCCAGTCAGCTGCCGGGTTCGCTGCCCGCGAACACCCCGAGGTCA
CCAGCCGCCAGCTACTTTGCCCGCGCTCCACGCCCTCCCTGCCCGCCCTGCCACCGGGACCGTGC
CGCGAGGCCAGCTACTTTGCCCGCGCTCCACGCCCTCCCTGCCCGCCCTGCCACCGGGACCGTGC
TCCAGCTCCACTCGTAGTCCCCGACTCCGCCAGCCCTGCCGTAGCGCGCTTCCCGTCCGGTCCC
GGTGGGAACCGCGTCCGCCCGCCCGCACCACTCGACGTTCCGCTGCCCGCTTCTGCACCCCTGG
AGCGCCGCGCTGGCTGCCAGCTCAAGTCGAAAGTTGCTCGGAAGTGCACGTCTTACGTGTC
AACAGAACGATGCCCTCCACAGAGATCAACCGTGTACATTGAAGATCTGCCCCAGGGTCTAC
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTCAAAGATGTTG
CAAGCTGCTTTGCTTACGTTACAAGAGTTGATGAATTCTCAAAGAAACTACTTG
AATGATATGTTGTAAGACATATGCCATTATACATGCAAATTCTGAGCTTAAAGATCT
AAACGTTACTACGTGGTGGAAATGTGAACCTGAAAGAAATGCTAAATGACTCTGG
TTCCGCCTGGTGAACCTCCAGTACCACTTACAGATGAGTATCTGGAATG
CCCTCGGAGATGTCCTCGCAAATTGAAGCTCCAGGTTACTCGTCTGGT
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCCGTGGTAAACCC
ATGATCTACTGCTCCACTGCCGGGTCTCGTACTGTGAAGCCATG
TGTTGGCCAACCAAGGGATCTGATTGAAACAATT
GAGGGTCTTCACATTGATCGGTATGGATCCCATCGATGTGA
AATAGTGTCAAGTGTCTCAGAGGTTTCCAGGGATGTGG
TCCATCTGAAAGTGCCTTCAGTGCTCGCTCAGACCACAT
AGTTGGACCGACTGTTACTGTGCAAGGAGAACTG
CTGTTGCAGTGACAGGAATGGATTAGCAACCCAG
ATACTGATCCTCGCAAATCATGGCTTCTCGAGTG
GACTCTTGATATCGTGAAGAGTGTGGAGAAGG
TTTGACTACAATGCCACTGACCATGCTGGAAAG
CTATAGTTAGTGTGCATTGATTTATC
TGGACAAACAATGTACAGTTTACTATGTGG
AGGAAAAGGGACTGTGCATTGAGTTGGTCTGCT
CTATAGTTAGTGTGCATTGATTTATC
TCTGAAATTTAAAGTGTACAGAAGCAG
App_ID=10063515

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKCSEVRRLYVSKGFNKNDA PLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNCNLQAVFASRYKKFDEF F KELLENAEKS LNDMFVKT YGH
LYMQNSELFKDLFVELKRYYVVGNVNL EMLNDFWARLLERMFR LVNSQYHFTDEY LECVSKY TE
QLKPF GDVPRK LKLQVTRAFVAARTFAQGLAVAGDV SKVSVNPTAQCTHALLKMIYCSHC RGL
VTVKPCNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDP IDV KISDAIMN
MQD NSVQVSOKV FQGC GPPKPLPAGRISRSI SEAFSARFRPHHPEERPTTAAGTSLDR LVT DVK
EKLQAKKF WSSLPSNV CND ERMAAGNGNE DDCWNGKGKS RYLF AVTGNGLANQGN NPEVQV DTS
KPDILILRQIMALRVMTSKMKNAYNGNDV DFFD ISDESSGE GSGSGCEYQQCPSEFDYNA TDHAG
KSANEKADSAGVRP GAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glycans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCCATTATATTCTTCAAGCAACT
TACAGCTGACCGACAGTTGCGATGAAAGTTCTAATCTTCCCTCCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTCTAGGAGATGGCTCCAGGAAGGCGGCAAGAATGTGAGTCAAAGATGGTTCCCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGCTGGGCTGCCAAAGAAGCAGTGCCCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACACAGAAAGCCAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAAATGTCAGCTAAGAAGCTTGCTCTGCCCTTGTAGGAGCTCTG
AGCGCCCACCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCGACACTC
TTCTTCTCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGTTGCTCTCTAGTGTCTCTCGTCAGTCTAGGCC
GTGCCCTCCCTTACCCAGGCTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCC
AGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCATAAAATTTTT
AAATGTCaaaaaaaaaaaaaaaaaaaaaa

FIGURE 26

MKVLISSLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPCKDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGGCCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTCCCCAGCGGAAGCACAGCTAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACCTCCTGCAGCTGCTGGTCTGCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCGTGCAAAAGCTACTTCCCACCTGA
TGGCGTCTGACTCCAAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTTCAGCCAG
ATAAAGGGCTTACAGGAGCCTCCGGAAAGTGGCCCTACTGGAGCTGGCTGCGGAACCGGAGC
CAACTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCATGGATGTGGTCTGCACTCTGGCTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTG
TCTTTTCTGGGAGCATGGCAGAACCATATGGAAGCTGGCCTCATGTGGCAGCAAGTTTC
GAGCCCACCTGAAACACATTGGGATGGCTGCTGCCCTACCAAGAGAACCTGGAAGGATCTGA
GAACGCCAGTTCTCCGAATCCAATGGAACGACAGCCCCCTCCCTGAAGTGGTACCTGTTG
GGCCCCACATCATGGAAAGGCTGTCAAACAAATTTCCAAGCTCCAAGGCACACTATTGCTCC
TTCCCCAGCCTCAATTAGAACAGCCACCCACCAGCCTATCTATCTTCAACTGAGAGGGACTA
GCAGAATGAGAGAAGACATTACGTACCACTACTAGTCCCTCTCCCCAACCTCTGCCAGGG
AATCTCTAACTCAATCCGCCCTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGATCCTCAACTGCAAGTTCTGGACTAGTCTCCAACTGTTGCCTC
CCAATGTTGTCCTTCCCTCGTTCCATGGTAAAGCTCCTCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTCTAGGAACGGTCACAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTGAGCTGGGGCACCAAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATTTTTAATAAATAGACGAA
ACCAAG

MDILVPLIQLLVLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL

FIGURE 28

MDILVPLIQLLVLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDDVVVCTLVLCVQS PRKVLQEVRRVLRPGGVLFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWPVGPHIMGKAVKQSFPSSKALICSFPSL
QLEQATHQPIYIPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTGCCTATCCACCTCCCCAAGCCCCTTAACCTATGCTGCTGCTAACGCTGCTGCTGCTGCTGCTGCTAAAGGCTCATGCTGGAGTGGGACTGGTCGGTGCCAGAAAGTCTCTCTG
CCACTGACGCCCCCATCAGGGATTGGCCTTCTCCCTTGTGTCTCCTGCCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAACAGAGGGTAGTGGTGGCTAGGGGGCTGCCTTATTAAA
GTGGTTGTTATGATTCTTACTAATTATAACAAAGATATTAGGCCCTGTTCATTAAGAAATT
GTTCCCTCCCTGTGTTCAATGTTGAAAGATTGTTCTGTGAAATATGCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCCTTATTTAACCGATGTGTCCTCAAACACCTGAGTGCTACTCCCT
ATTGCATCTGTTGATAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGGAAATCATGTCGG
GAAGAGATAACATCCTGGGCTGTGTATCCTCGCATTAGCCTTGCTTGGCCATGATGTTAC
TTCAGATTCACTACCCACCTCTGGTTCACATTTCATGGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAATATGAAGTGCCTGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCCTCCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGGCTCTGGTGGCTGTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCAGGCAAGTGAATATAAGCCCCTTCGGGCATTGGTACATGTGGCGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCACTCCCTGGTGCCAGCAAATGACTATAGCTGGGCAG
TGGTTACTTGTATTCACAGAAGTAAAATGATCCTCCTGATCATCCATCCTTGTCTCTC
TCCATTCTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTATTTATCTGTGGTGAG
GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTCATTG
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCGAAATGCATATACTACAACGTCTATTGAGCAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTGTCCAAGAAACTCAAGTCACCTTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTGTTACTGTTGGAGACTCATGGCTTTACTAC
AATCGGGCATTCCAGGTGTTGGCAGTCCTCTGTTATTGGTAGCTTTGCCTACTTAGTGC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCCCTGTGTTGCTGTTGATC
TGGAAACAAATGATGGATGTCAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTGTA
AAAAGGAGCAACAAATTAAACATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACCTCCAGGCCATTGTGAGATAGATACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTCTAAGAGCCATTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTT
TTAAAAGACCTAATAAACCTATTCTTCTCAAAA

MSGRTDILGLCILALALSLAMMFTFRFITTLLVHIFISLVLGLLFVCGVLWWLYYDYNLDSIE

LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISAPFLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMVMQNALKEQQHG
ALSRYLFRCYCFCWCLDKYLLHQNQAYTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPTLMLQLLAVTVAVA
TPPTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVLPKEELSTSLEPVSLGDPNCKIDLSSLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVVMDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPQPLVKRVCDTDRACSKTCLNSADIGFVIDGSSVGTGNFRTVLQFVTN
LTKEFEISDTDRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALQL
FKKSXPKNRKLMILITDGRSYDDVRIPAMAHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTAGGAGGTGGCTGCGTGTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTCTGTTTCAGAGTAGTTACAACAGATCTGAGTGTAAATTAAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCATCTGAAATTCCACAGTTCTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTCCCGCTGGCTCTACCGTGGTGCCTCCGACTACTCACCCCGAGTGTA
AAGAACCTTCGGCTCGCGTGCCTCTGAGCTGCTGTGGATGGCCTCGGCTCTGGACTGTCCCTGGAGTA
GGATGTCACTGAGATCCCTCAATGGAGCCCTGCTGTCACTCCTGAGTTCTTGTGATGTGGTAC
CTCAGCCCTCCCCACTACAATGTGATAGAACCGCGTAACGGATGTACTCTATGAGTATGAGCCGATT
CAGACAAGACTTCACTTCACACTTCGAGAGCATTCAAACACTGCTCTCATCAAATTCATGGTCAATT
TGGTGACCTCCCAACCCCTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGTGAAAAAAAGTCT
TGGTGGGATATGAGGTTCTTACATTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
ATTGTCCTTAGAGGATGAACACCTCTTATGGGACATAATCCGACAAGATTTCAGACATATAATA
ACCTGACCTTGAAAACCATTATGGCATTCAAGGTGGTAACTGAGTTGCCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTCATCAACTGGCAATTAGTGAAGTATCTTTAAACCTAAACCACTCAGA
GAAGTTTCACAGGTATCCTCTAATTGATAATTATCCTATAGAGGATTACCAAAAAACCCATATT
CTTACCAAGGAGTATCCTTCAGGTGGTCCACTGCAAGGTTAGAGGAGATCATCAGTT
TTGGTGCCAAGGATCTATGAAATGATGGTCACGTAACAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATAAAAGTGAACATTCAATTCCAGAACAGACACAAATCTTCTTATATAGAATCC
ATTTGGATGTCTGCACTGAGACGTGTGATTGAGCCCATGGCTTTCTTCAAGGAGATCATCAGTT
TGGCAGGTCACTGCAAGGAACACCACATGCCATTATTAACTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACTGTGAAAGTGTAAATAAAAGTAGGTACTGTGAAAATTCAAGGGAGGTCACTGTGCTGGCTT
ACACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGTTACACTTGTGATTATTAGTCAGG
CCCTCAAAGATGATATGTGAGGAATTAAATATAAAGGAATTGGAGGTTTTGCTAAAGAAATTAAAGG
ACCAAACAAATTGGACATGTCAATTCTGTAGACTAGAATTCTTAAAGGGTGTACTGAGTTAAAGCTCA
CTAGGCTGAAAAACAAAACATGTAGAGTTTTTATTGAACATGTAGTCACCTGAAGGTTTGTA
TATCTTATGTGGATTACCAATTAAAATATGTAGTTCTGTGTCAAAAAAACTTCTTCACTGAAGTTATA
CTGAACAAAATTTCACCTGTTGGCATTTATAAAGTACTTCAGATGTTGCAAGATGTTGAGTATT
ATTATTTAAATTACTTCACATTGTGTTAAATGTTGACGATTCAATACAAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTTCAGTTACTTAACAGTACAGTTATTATTGATAACATCACTCCA
TTAATGTAAGTCATAGGTCAATTGCAATTGCAAGTACTCTGGACTTTGTTAAATATTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

.....L.....P.....R.....S.....L.....K.....W.....S.....L.....L.....S.....F.....F.....V.....M.....W.....Y.....L.....P.....H.....Y.....N.....V.....I.....R.....V.....N.....W.....M.....Y.....F.....Y.....E.....P.....I.....Y.....R.....Q.....D.....F.....H.....

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNC SHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLGQEAEKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIIHIEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYT VINEACPGAEWNIMCRECCYEYDQIECVC PGKREVVGYT
I PCPCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGT LDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRCVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCCQNGEWSGKQPICIKACREP KISDLVRRRVLPMQVQSRETPHLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCI PICGKIENITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAFLVCSGALVNERTVVVAAHCVTDLGKVTMIKTADL
KVVLGKFYRDDRDEKTIQSLQISAIILHPNYDPILLDADIAILKLLDKARI STRVQPICLAASR
DLSTSFOQESHITVAGWNVLADVRSPGFKN DTLRGVVSVDSSLCEEQHEDHGI PVSVDNMFC
SWEPTAPS DICTAETGGIAAVSFPGRASPEPRWHLMG LVSWSYD KTC SHRLSTAFTKVL PFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

SUB domain proteins profile

amino acids 150-167

FIGURE 39

GGTTCCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTCTACGGGGCCGTGATTTATTAACTGGCTTAATC
TGAAGGTTCTCAGTCAAATTCTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTTAAAGGAGCTGGCTGG
TTGGGCCCTGTAGCTGACAGAAGGTGGCCAGGGAGAATGAGGCCTTGTGAGC
TGGCTTGCCTGGCTCAGTCTGTACTACATTGACAATGTGGCAACCTGCACTTCTGTATTCAGACTCTGTA
AAGGTGCTCCCACTACGGCCTGACAAAGATGGAAAGGGGGCCTCACAAAGATGGCTGTCCAGCGGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCAGGAGTTCTGCAGCTGCCCACCAATCCTTAATGACAGACCGACCCGTGGCTAGACA
ACCCCTGCCTACGTGTCCTCGGCAGGACGGCCAGCCCAATCCAGGGACTCTGGCCGGCAACCGAACT
GGGCACGGCCTTGGAGATCCACTATTAGAACGAATCCTTAAAAATCGAGCTTTGGAGTGTCTTCGAAA
GGACAAAGGGCCGGAGGTGCAGTTGCCAACCAATCCGCACGGCCAGGGAAATCTGAAAACCCACGTCCCTGAAG
TCTTTCCAAGGTTGTACCCACCTGTATCCAGATGGTGAAAATACCAGCATCAAGATCGAGTAGTCCAGGTGAAA
GCCTCTCTTATTAGGCGTGGGGAGGTACGGAAACCCACCTGGCCATATCATTATCCAACACAACATTTATGTGTGG
TGATCGCCAGACGGCCGGCTACGTGCAGGACATCATTCAAAGGTCAACGGATGGACATCGCAATGTCCCTGA
ACAAACACGTCTGTGCGTCTCTGCGGCAGGCCCTGCAGGGGTGTGGCTGACTGTGTGAACGAAGTTCCGACA
GCAGGACAATGGACAGGGCCGGATGCCTACGAACCCGGAGATGACGCTTTCAGTGTGATTCTCAACAAAGTAGCCC
CCGGAGGGAGCAGTTTGGAAAAACGTGGCGCAAGGTGGATGCGGCTGGGTTTTCATTCAAGTGTGGATGGCC
GTCTGGCAATTCGACTGGTCAGCTGTGGGAAGATGACGCTGTGTGGATGCCCATGATCATGGCAATGGCAAA
GCCCCAGAAGTCCGGCTCATCGTATTCGGATCCACCTGGCCAGGGGTGTGGCCAGGGGTCCGACC
GGAGCCCTGACATCTTCAGGAAGCCCGGCTGGAACAGCCAATGGCAGGCTGGCCCCAGGGGAAGGACAACAA
CTCCCAAGCCCCCTCCCATCCAAATTACGTTCATGAGAAGGTGGTAAAATCCAAAGCCCCGTGAATCTCTCG
GCATGACCGTCGCAGGGGGGAGCATCCAATGAGAAGGTGGATTTGCCTATCTATGTCACTAGGTGTGGAGCCGGAGGAGG
TCATAAGCAGAGATGGAAGAAAACAGGTGACATTTTGTGAATGTGGATGGGTCGAACTGACAAGGGGTCACC
GGAGGTGGCAGTGGATTTGAAAAGAAACATCATCCTGTGATAGTACTCAAGGTTTGGAGGTCAAAAGGTATGAGC
CCCAGGAAAGACGTCGCAGCCCAGCGACCCCTGGACCTCAACCAACATGGCCCCACCCAGGTACGTGGCCCCATCT
GGGTCAGTGTGGGTGAAATTACACGGGTGTTAACTGTAAAGATTTGTATTAACGAAAACACGTGGAAAGTC
TGGCTCTGTCATTGTAGGGGTTTATGAGAAAATCAATGGAAACAACCTTTTTCATCAAAATCCATTTGTGAAGGA
CACCACAGCACAATGTGGAAATTAGAAGATGTGGGTATTTCTCTGTCAATGGTAAGGTACATCAGGAAATGA
TACATGTGTGGGTGGCAACACTGTGTGAAAAGACCTAAAGGAAAATTACTCAACTATTTGTTTCTGGCCGTGGCACT
TTTTATGAATCAATGTGGGTAGGGAAACAGAAATCAAAATAGGTTAAGGTAAACACATATTTATTC
TGTCAGTTTTTATTTAAAGAAGAATCAATGTAAAGGTGGAAAGATGTGTGAATCAATGTTAAGAAAGCCAGTT
ACACCTCAAAATATGTCAAAAAAAATTTAAACTAACGTGGAAACCTCAAAATATGTGTGAATCAACCCACGTGAAA
AAACATGTTTTTATTTCAAAAAAGCCCTAAACACCAAAATGTGAATTTGTATACCCACGTGAAA
CAAGGTATAAATTTAAATGTGTATATGTCAGTGAAGGTCTGTCCAAGGGTACTTATGGCCATTTAATTACAGCT
AAAAATTTTTAAATGTCAGTGGGTGAAAACGTGTGGATTTTCATCAAAACGAAAATATTTTCAGAAGTAAAA

FIGURE 40

MKALLLIVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGASC LTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESL SIRLV
GGSETPLVHIIQHIYRDGVIARDGRLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTV
REQKFRSRNNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI FNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVSRQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTGDILLNVDGVELTEVSREAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFKIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTCAGTTGTCAAGTCGAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCCATAAACTTCCGATTACTGTTGCTGTTGACTTTGCTGACAGTGGTTGGGTGGC
CACCACTAAGAGATTCTAAAGCAAAGGAGTCATGGCTAATTTCC
ATAAGACCTCATTGGGAAAGGGAAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACCT
GACAACGTGCTCTGTGTCCTTACCTCAGAGGCCAGAGCAAGCTCATTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGTATGCCCTCAGGAATGTAAG
CTTACAGAGGGTCGCCATCCTCGTCCCCACCGGAACAGAGAGAAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTATCCACCAGGCTGAAGG
TAAAAACTTAATCGAGCCAAACTCTGAATGTGGCTATCTAGAACGCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTAACTTACAAGTGTGAGGAG
CATCCCAAGCAGCTGGTGGTGGCAGGAACAGCACTGGTACAGGTTACGTTACAGTGGATATTTGG
GGGTGTTACTGCCCTAACAGCAGAGGCACTTTCAAGGTGAATGGATTCTAACAACTACTGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAAGAATGAAAATTCCCGCCCTG
CCTGAAGTGGTAAATATAACATGGCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTAACCCAAGTGTACAGACTGGAGAACAGATGGTTGAGTAGTTGTTCTATAAAT
TAGTATCTGTGAAACACAATCTTATATCAACATCACAGTGGATTCTGGTTGGCATGACCC
TGGATCTTGGTGTGATGTTGAAAGAACTGATTCTTGTGCAATAATTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCTGGTGTGAGACTATAAAACAGTTGTAACAAGACAGCTTCTAGTCATTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTTGAAGGAGATTATTTAAAGTGAAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCAGAGAGAACAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAACTGTTATTCTTATCCTGTACAATCATGTAAGTGGTGGTGTAGGT
GAGAAGGCGTCCACAAAGAGGGAGAAAAGGCGACGAATCAGGACACAGTGAACGGGAATGAAGA
GGTAGCAGGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGATAGC
CTTCAGGGAGGACCTGCCAGGTATGCCCTCCAGTGTGATGCCACAGAGAATACATTCTCTATTAGT
TTTAAAGAGTTTGTAAAATGATTTGTACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAACACCTCTAGTAAAATGTGAAAAGCAAA

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN

EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFIHHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGVTALSREQFFKVNNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCAGATCTGGG
CCGCTTGCCTCTGCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTCCCACAACAG
ACGGGACAACCTGCAGAGCTGCAACCCAGGACAGAGCTGGAGGCCAGCTGGATGCCAT
GTTCCAGAGGCGAAGGAGGCAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGATGTGCTGCAAGACGTTAGAACCTACCTGCCCTGCCCTGCCCTCCC
TTCCCTTATTTATTCCTGCTGCCCTGAGAACATAGGTCTTGGAAATAAAATGGCTGGTTCTTTGTTT
TCCAA
AAAAA

FIGURE 44

MALSSQIWAACLLLLL~~ASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI~~
CIFCCGCCHRSKCGMCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTCCCCCTGAAGTCAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTTGTCAACATACAGCCAGAAGGGGGCACTATCATAGTGA
AAGAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAA
ACTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAG
CCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGAGAAAGTGTATGACCTCATCTGCGTTGCCAGGAACCCCTGTCAGCAGAA
ACT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCTGCTCAGTCTTTGACTGGGCTATTCTT
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTGTC
GGG
AAACTCCTAACATATGCCCTATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAACTCCTAACAGAAGATCCAGCAAATACGGTTACTCCACTGTGAAATACGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLKSVKQVDSIVWTFTNTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSMVLCLLVPLLLSSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCCTCAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCCTAAACCCCCATCTCTGCTTGAGTGGTGGTTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCCAGCAACAACAATGTCCTGACAGCAAGAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCCATCCAGGCCTCTTAAAGGTCTCTCATGTGAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTCAA
CTTGCAGTGGTTTCATGACTCTTGTCACCTCCTACTGGTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTAGGTCTATTGCTGTTGAAATTCTGGAGGTCTGTTGGGCT
CAGTCAGATAGTCATCGGTTCCCTGGCTGCTGTGGAGTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLGVVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCGCCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACGTGTA
CATGGAGAGAGTGACCCCTGCCCTTCTCCTACTGGCAGGCCGTGACTGCCTTGGAAAGCCAATGACC
CATTGCCAATAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCAGGGAGGGCTCCTGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCGTACCTGAGAAGGCCATCCACTCATCACCCAGGCTCTGCCA
CTACTTGCT**GACAGGACTGCCCTCAGGGATGCCGTGAGCTAACACTGGCCCCAGCACC**
TCCTCCCCCTGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTGCCACCACCCCTCA

21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

FIGURE 50

MERVTLALLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLIAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCCCTCGCTGGCCCTCTGCCTGGCAGTGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCAGTGGACAAATATTGGGGAGGCCCTGGACATGGCCTGGAGACGCCCTGA
GCGAAGGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGTTGGC
AGCAGATGCTTGGCAACAGGGTCGGGAAGCAGCCATGCTCTGGAAACACTGGCACGAGA
TTGGCAGACAGCAGAAGATGTCATTGACACGGCAGATGCTGTCCGGCTCTGGCAGGG
GTGCCCTGGCCACAGTGGTGGCTGGAAACTTCTGGAGGCATGGCATCTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCCG
GAAACTCAGCAGGCAGCTTGGAAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGGAGCAGTGGCAGTGGCAGCAATGGTGACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCAGCAGTGG
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGAGCAGGGAGGAAATGGA
CATAAACCCGGGTGTGAAAAGCCAGGGAAATGAAGCCCGGGAGCAGGGAAATCTGGGATTCAAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGCAATGCCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGAGTGGGAGGAGGTGACGCTGTT
GGTGGAGTCATACTGTGAACCTCTGAGACGTCTCTGGGATGTTAACCTTGACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAGGACAGAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGAGGCCACACTCCCTTAA
AACACCACCCCTCATCACTAATCTCAGGCCCTGGCCTTGAATAACCTTAGCTGCCACAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFAADALGNRVGFAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGHGIFGSQGGLGGQGQGNPGGLGTPWVHGYPGNSAGSGMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPSGSGGGSSNSGGSGSQSGSSGSGSNGDNNNGSSSGS
SSGSSSGSSSGSSGGSSGSSGSSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNYRGQGSSWGSGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGGCTGGGCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGGTTGTGGGCTCCTGGCTACTCGCCCGCATCCTGGCTGGACCTATGCCTTCTATAACAACTGCCGCCGGCTCCAGTGTTCACAGAGGAGGGCTTGAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTACGGTATGGCTGGGCTCCATCATCCCCTTCATCGTTTATGCCACCTGACACCATCCGGTCTATCACCAATGCCCTCAGCTGCCATTGCAACCCAAGGATAATCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGGCCACCGTCGGATGCTGACGCCGCCCTCCATTCAACATCCTGAAGTCCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGGCCTCAGTCAGTCGCTGGACATGTTGAGCACATCAGGCTCATGACCTGGACAGTCTACAGAAATGCATCTCAGTTGAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTCTGATTACCTCTCCATGACGGCGGCCTCCACAGGGCCTGCCGCCCTGGTGCATGACTCACAGACGCTGTCATCCGGGAGCGGGCTCGCACCCCTCCCCACTCAGGGTATTGATGATTTTCAAAGACAAAGCCAAGTCCAAGACTTGGATTCATTGATGTGCTTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTGGAGGCCATGACACCACGGCCAGTGGCCTCTGGTCCGTACAACCTTGCGAGGCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGCAAGAGCTCTGAAGGACCGCGATCTAAAGAGATTGAATGGACGACCTGGCCAGCTGCCCTGCATCCTCGATGCTGCCCTCCTGGGCTCTGACAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCTCATCGATATTATAGGGTCCATCACAAACCAACTGTGTGCCGGATCTGAGGTCTACGACCCTTCCGTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTATTCTTCTCCGAGGGCCAGGAACACTGCATGGGCTGGGAGATGAAAGTGGCTCTGGCGTTGATGCTGCTGCACTCCGGTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGCGCCGAGGGCGGGCTTGCTGCCAGGGCTGAATGTAGGCTTGCAGGTGACTTCTGACCATCCACCTGTTTTGAGATTGTCATGAATAACGGTGTGCAAA

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTCTGTTGACTTACCCGTGTTCTGCCT
TTTGGTGGCATAACAAAGGGACTGCACTTACCTTCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTTGC
ATAATTCATGGCCAGTTTATGAAGCTTGGAAAGGCACTATGGACAGAAGCTGGTGGACAGTTT
GTAACATATCTCGAAACCTCTGTCTTACAGACATGTGCCTTTATCTGCAGCAATGTGTTGCTT
GTGATTGAAACATTGAGGGTTACTTTGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTGTGGAGTGGAAATCTTCCTCATGTACCTGTTCC
TCTGGATGTTGTCCTGAATTCCCATGAATACAAACCTATTGCAACAGCAAAAAAAAAAAAA
AA

MGPKQLKRMFEPTRLIATIMVLLCFAVWGNKGLALIFCILQSLALTWYSLSFIPFAR

FIGURE 56

MGPKQLKRMFEPTRLIATIMVLLCFAVWGNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGCAGTGGAGCACCCAGCAGGCCCAAATGCTGTCGTGCCCTG
TACGTGCCGTATCGGGAAAGCCAGACCGAGTTCCAGTACTTGTAGTCAGAAGGGCTCCCTGCCAGTCAGAAGCTC
ATTTCAAGCTCAGTGTCTTATCCCCCTCCAGGAATTCTCCACCTACGGCAGTGGAAAGCAGAAATTGTAACAGG
GGAGATAAGGACCTTATGGGAGCTAGACATTCTGGACAGTGGAAAGTATTGTCATTATCTCAAAGATCATGAGAAGAGCTGAGG
CTGGTGTAAGATTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCCGGACTTG
GGAGTCAGAATATCTGAACAGCAGGCCAAGAAAATTCAGAAGACATGGATAAAACGGCAGATGCCATTGACCCATGAGCTGG
AACAGTGGAGACGACTCACCTCTCACCCTCAGGGAAACATCAGGAGATCATCTCTAGTGGAGAGAGGGCAGACGGGATGTGGTGAGA
ATCTTGATGTGGGAGATAAGCTAAGCTCCCGAGATGAGTTCAAGCTGGAGGAGAGGGCAGACGGGATGTGGTGAGA
CACCTGGTGGCAGGAGGGGGCAGGGCGTATCCAGAACCTGACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTTCACTGCCCTCCCGAGCAACAACTGAGGATCTGGTGGCTCACTCAGATGATTGAGAGGGGGCAGG
TCACTCTGGGGGGCAATGGCATCAACTCTCTCAAAATTGCCCCGAATCAGGCAATCAAAATTCTGGGCTATGAGCAG
ATCAAGGCCCTGGTGTGAGTCAAGCAGGAGACTCTGAGGATTCAAGGAGGGCTGTGCAGGGGCTTGGCAGGGGCC
ATGCCCAAGAGCAGCATCTACCAATGGAGGCTCTGAAGACCCGGATGGCGCTGCCAGAGCAGGGACTACTCAGGA
ATGCTGGACTCGGCCAGGAGGAGCTCTGGCCAGAGGAGGGGTGGCCCTTCAAAAGGCTATGTCCCAAACATGCTG
GGCATCATCCCTATGGGGCATGACCTGAGCTACAGACGCTCAAGAATGCCATGGCTGCCAGACTATGCAGT
AACAGCGGCCGACCCGGCTGGTGTGTGCTCTGGGCTTGGCACCATGTCCAGTACCTGTGGCAGCTGCCAGCTAC
CCCTGGCCCTAGTCAGGACCCGGATGCAGGGCAAGCCTCTATTGAGGGCCTCCGGAGGTGACCATGAGCAGCCTC
TTCAACATATCTGCGACCCGGGGCCCTTCGGGCTGTAGCAGGGGCTGCCCAAATCTGAGGATCTCA
GCTGTGAGCATCAGCTACGGTGTCTACAGAGAATCTGAGATCACCTGGGGCTGCACTCGCGGTACGGGGGAGGG
CGCCGGCGACTGGACTCCGTATCTGGCCCGAGCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACT
AAGCTGTCTCGAGCCAAGCTGTGAAACCCTAGACGCAGCCCGAGGGCTGGGGAGAGCTGCCAGGCCAGGGCTT
GTCTGCTGACCCAGCAGACCCCTCTGGTGGTCCACGGAGACAGGCAACAGGCAATCTTCTGGGCTCAGGGTCAGAG
CTCCGGCTCACATGTGAAGGACAGGACATTCTGCAGCTGCCAAATGTGAGCTTGGACCTGGAGGGCCCT
TAGTTCTTCAATTACCCCTGCAGCCAGCTGTGGGCCAGGGCCCTGCCCTCTGGCTGTGCGTGCATCTCCCTGTG
CCTCTTGCTGCCTGCCGTCTGCTGCTGAGGTAAGGTGGAGGGCTACGGCCACATCCACCCACCCCTGTCCAATCC
ATAATCCCATGTAAGGAGGTGAGGTCACTGGCTCCACGGGCTGAGTCCCAACCTACAGCATTGACGCCAACACTTGG
TGTGAAGGAGAGGAAAGGATCTGGCTTGTGCTACTGGCCATCTGGGCTCTGGGCTGCAGCCCTGTGCTGGGGCT
CTTGGGAGTGAGGGGCTGGGCTGGCTGGCTGGCTGCAGAGGAGGCAAGTGTGGGCTCATGGTGTCTGAGCT
GCCCTGGACCCCTGCAGGATGGGCCACCCACTGAGAAAACTCATGCTCCCACTGTGCAGAGGGCAGTGGAGCA
CCATGTTTGAGGGCGAGGGCAGAGGGCTTGTGTTCTGGGGAGGGAGGGAAAGGGTGTGGAGGGCTTAATTATGG
ACTGTGGGAAAGGGTTTGTGCTCAAGAGGACAAGGCCGACAAATGAGCAGTCTGTGCTTCCAGAGGAGACGAGG
GAGCAGGAGCTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGTTCCTGTCCAACCCAGCAGGGCGCAGC
GGGACCAGCCCAACATTCACCTGTCCACTGCAGCTTGTGAACTATTTTTGATGGAGAGGAGTTGCT
AACTATTTTTGATGGTTTAATTATAGCTTGTCAGTTTCAAGTTCATTTTTTTGATGGTCAGGGTT
GATTGTCACCTTCCCAAGGGCCGGCACTGGGGATGGGGAGGGAGGAGGGGGCCCTGGGCCGTGCAGCTCACATCT
GTCCAGAGAAATTCTTGGACTGGAGGAGAAAGCAGGCCAGAGGCAGAGGCCAGGCCAGGGCTCCCTCCCTGGCAG
GTTGGGGAGGGCTTGGCCCAAGCCTAGGTTCAGGGGTTGACTGGGGGCTGGAGAGGAGGGAGGAAACCTCAAT
AACCTGAGGTTGGAACTGGTCAGGGTCTGGGCTGGAGGGTTTCTTATTTCTACGTCTTGTCAAGGCG
TGAGGTGCTCTCACTGTGAGTTTGTGCTGGGGGGGGCTGGAGGAGGGGGCTGGCTCGTCCCTCCCAG
CTTCTGCTGCCCTGTCTAACATGCCGCCAGCTGGCAGCTCACGGTGTGCAGCTTCCATTCCACCAGAAATGACCTGA
TGAGGAATCTTCAATAGGATGCAGAAATGTCAGAAAATTGTTATATGAACAATTACTACTGGAGTGTGCAG
CAAATTAAGGAAAATTGGACGCTTAGAGTGTCAGTTAAAGGAGCCTTCATAAAATGTGTTCAAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFI
PSQEFSTYRQWKQKIVQAGDKDLDG
QLDFFEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMKDNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTV
PDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINV
LKIAPESAIIK
FMAYEQIKRLVGSQETLRIHERLVAGSLAGAIAQSSIYP
MEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETL
KNAWLQHYAVNSADPGVFVLLACGTMSS
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGA
FGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATAACGCTGGAACCTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTGAACCTGACATCAAACCTTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTGGCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGGCTGGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTTTGCGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTCTAAAGGCAAGGGATGCTAACCTGAGTATAAAACTGGAGCCTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCCTGAGGCTCCCGATGGTCCCC
CAGCCCCACAGTGGCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACCTGAGAATGTGACCATGAAGGTTGTGCTGTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAAATCGGAGATCAAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTG
CTCTCTTCTTGCCATCAGCTGGCACTTCTGCCTCTAGCCTTACCTGATGCTAAAT
GTGCCCTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACCAGATATGACCTAGTTATATTCTGGGAGGAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGGCCAAAGCAGAAGGCTCAATATGAACAAGATAAT
CTATCTCAAAGACATATTAGAAGTTGGGAAAATAATTGACATGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGAGTGAGAGGACAGGATAGTGCATGTTCTGTCTGAATTAGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCTGAAAGTCTATCCAAACATATCCACATCTTATATTCCAC
AAATTAAAGCTGTAGTATGTAACCTAAGACGCTGCTAAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGCAAATGATTCACTTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCAGTTATAAATAAACTGAGCACCTCTTTAAACAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

.....

FIGURE 6o

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYIMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCAATGGCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCCTCGGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCCTGGTGGTGGTTATGGGGTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCTGGGATGTTCCCCTGGAACCTCCAGGAGGACCATATG
GCGGTGAGCTCCGGGGGGCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCAGCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCAGCCCTGTGAAATTCACTCCAGCAGTGGAGAACCTCTT
CCAGCAGTATGACGGGACCGCTGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGG
AAATGGGCTACAACCTGAGCCCCCAGTTCAACCCAGCTCTGGCTCCCGCTACTGCCACGCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTGAGGACTTCGTCA
CCATGACAGCTCTCGGATGCTATGACCAACCATCTGTGGAGAGTGGAGTGCACCAGGACCTT
TCCTGGCTTCTAGAGTGGAGAGAAGTGTGGACATCTCTTCTTCTGTCCTCTAGAAGAAC
ATTCTCCCTTGCTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTGCATCATAGCCACCA
AAATAGTGGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCCTGATGCCAGTGGTGGAGTCTGCATGGCTGTTACC
GTTAGTACCTGTTCCCTCACCAAGGCATCCTGTCACAGGCCATTTCCTCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACAGTGGCTTGGATTCTGCCACACCCATAAAAT
CCTTGGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTTGCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCAATTTC
ATTGGGGCAAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTPGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSAWKFICQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGG**A**TGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTGCCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTCCAAAACCTCGGGACC
AACTTGCCCTCAGCTCGGACAACCTCCTCCACTGGCCCTCTAATCTGAACATCCGCAGCCC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGG
AGCGCTGCCTGAAGAACTCTTACCTCTCCAGTGCCTGCCAGGCTCTCACCTGAGGCTTCACTCCT
TGCCTGGGGAGTCTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGAGCCGGGGAAAATCCTTCCAACG
CCCTCCCTGGCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCCTGAATCCAGTG
TGTCCCTGGGGAGGTGGAGGCCCTGGACTGGTTGGGAACGAGGCCATGCCACACCCTGAGGGA
ATCTGGGTATCAATAATCAACCCCCCAGGTACCAGCTGGGAAATATTAAATCGGTATCCAGGAG
CAGCTGGGAAATATTAAATCGGTATCCAGGAGGCAGCTGGGGAAATTAAATCGGTATCCAGGAG
GCAGCTGGGGAAATATTCACTATACCCAGGTATCAATAACCCATTCCCTGGAGTTCTCCGC
CCTCCTGGCTCTTGGAACATCCCAGCTGGCTTCCCTAATCCCTCAAGCCCTAGGTTGCAGTG
GGG**C**AGAGCACGATAGAGGGAAACCAACATTGGAGTTAGAGTCTGCTCCGCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCACCTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

FIGURE 64

MQGRVAGSCAPLGLLVCLHPLGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRMPMPHEGIWGINNQPPGTSGNINRYPGGS
WGNINRYPGGSGNINRYPGGSGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCTTGCTCTCTTGCACCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTCATGAGTCCAGCCTGGAA
TTGCTGAAAAGCTCTGCCCTCCTCCATCCCTCAGGGACCAGCGTCACCCCTCACCAGTC
AAGATCTAACACCATGTTGCTGCAACACATGACAGCCATTGAAAGCCTGTGTCTTCTGGCCC
GGGCTTGGCCGGGATGCAGGAGGCAGGCCCGACCTGTCTTCAGCAGGGCCCCACCTC
CTGAGTGGCAATAAAATTCGGTATGCTG

MGSGPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLHLPSGTSVTL

FIGURE 66

MGSGPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGCAGGTGCCCTGGCCGGAGATCGGGTAGGAGGGCGAGCGCGAGAACCCCC
TTCCTCGGCCTGCCAACCCGCCACCCAGCCCCATGGCGAACCCCGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCGAGCCTGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTGCCCTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGCTTCTCCCTCTGGCTGCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCAGCTCAGGGAGAACGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCCTCCTGCATCTGTCTCCCTCATTGCTGTGACCTTGGGAAA
GGCAGTGCCCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAGAACGGTACTTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATTTATATAAAATTAG
TAGTGAGATGTAaaaaaaaaaaaaaaaaaaaa

FIGURE 68

MANPGLGLLLALGLPFLLLARWGRAGWQIQTTSANENSTVLPSSSTSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFShAAEAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCAGGAATAACTAGAGAGGAACAATGGGTTATTCAGAGGTTTGTTCCTCTAGTTCTGCTGCTGCACAGA
TCAAATCTTCTCATTAAGCTAATAATAATGGCTTGAAGATATTGTCTTGTATAGCTCTAGTGTGCCAGAA
GATGAAAAAAATAATTGAACAAATAGAGGAATATGGTGACTACAGCTTCTACGTACCTGTTGAAGCCACAGAAAAAAAGA
TTTTTTCAAAAATGTATCTATATAATTCCCTGAGAATTGGAGAAAATCTCTAGTACAAAAGGCCAAACATGAA
AACCATAAACATCTGATTTATAGCTGACCCACTACACTCCAGGTAGAGATGAACCATACACCAAGCAGTTACAA
GAATGTTGGAGAGAAAGCGAATACATTCACTCACCCCTGACCTTCACTGGAAAAAAACAAATGAATATGGACCA
CCAGGCAAACACTTTGTCCATGAGTGGGCTCACCTCCGTGGAGGTGTTGATGACTACAATGAAGAGTCAGCCTTTC
TACCGTGCTAAGTCAGGAAATCGAACAGGTGTTCCGAGGTATCTCTGGTAGAAAATAGAGTTATAAGTGT
CAAGGAGGCAGCTGTCTTAGTGTAGCATGAGAATGTTCTACACAAAATGTATGGAAAAGATTGTCAATTCTT
CCTGATAAGTACAAAAGCATCCATAATGTTATGAAAGTATTGATTCTTGTGAATTGTAAACGAA
AAAACCCATAATCAAGAACCTCAACCCCTACAAAACATAAAAGTCAATTTCAGAACTACATGGGAGGTGATTAGCAAT
TCTGAGGATTTAAACCCATACCCATGGTGAACACCACCTCCACCTGCTTCTCATTTGTGAAGATCAGTCAA
AGAATTGTGCTTAGTTCTGATAAGTGTGAGCATGGGGGTAAGGACGCCCTAAATCGATGAATCAAGCAGCA
AAACATTTCTGCTGAGCTGTGAAAATGGATCCTGGTGGGGATGGTCACTTTGATAGTACTGCCACTATTGTA
ATAAGCTATCCAAAATAAAAGCAGTGTAGAAAGAACACACTCATGGCAGGAAITACCTACATCCTCTGGGAGGA
ACTTCCATCTGCTCGGAATTAAATGTCATTTCAGGTGATTGGAGAGCTACATCCAACTCGATGGATCCAGAAGTA
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTTCTGTATTGATGAAGTGAACAAAGTGGGCCATTGTCAT
TTTATTGCTTGGGAGAGCTGCTGTGAGGACTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
TCAGATGAGCTCAGAACATGGCTGATGCTTGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAG
TCCCTCAGCTCGAAGTAAGGAAACACTGAATAGTAAATGCTGGATGAACGACACTGTCATAATTGATAGTACA
GTGGGAAAGGACACGTTCTTCATCACATGAAACAGTCTGCCCTCCAGTATTCTCTGGGATCCCAGTGGAAAC
ATAATGGAAAATTTCACAGTGTGATGCAACTTCCAAAATGCCCTATCTCAGTATTCCAGGAACCTGCAAAAGGGGCACT
TGGGCATACAATCTCAAGCCAAGCGAACCCAGAAACATTAACCTACAGTAACTCTCGAGCAGCAAATTCTTCT
GTGCTCTCAATCACAGTGTGAAATGAGCTAAAGGAGCTTACAGTCTCCAGGCCATGTTACGAGAA
ATTCTACAAGGATATGTACCTGTCTGGAGCCAATGTGACTGCTTCTGATCACAGAATGGACATAACAGAAGTT
TTGGAACTTTGGATAATGGTGCAGGCGCTGATTCTTCAAGGATGATGGGACTCTACTCCAGGTATTTCAGCATA
ACAGAAAATGGCAGATATGCTTAAAGGCTCATGGAGGACAAACACTGCCAGCTAAATTACGGCCTCCA
CTGATAGAGCCGCTACATACCCAGCTGGTAGTGAACGGGAAATTGAACAAACCCGCAAGACCTGAAATTGAT
GAGGATACTCAGACCCACTTGGAGGATTTCAGCGAACAGCATCCGGAGGTGATTGTGATCACAGTCCCAAGC
CTTCCCTGGCTGACCAATACCCACCAAGTCAAATCACAGCTGATGCCACAGTCTGAGGATAAGGATTATTCTT
ACATGGACAGCAGCAGGAGATAATTGATGTTGAAAGTCAACCTTATATCATAAGAATAAGTGTCAAGTATTCTT
GATCTAAGAGACAGTTGATGATGCTTCAAGTAAATACTACTGATCTGTCACCAAGGAGGCCACTCAAGGAA
AGCTTGCATTTAACAGAAAATCTCAGAAGAAAATGCAACCCACATATTGCAATTAAAAGTATAGATAAAA
AGCAATTGACATCAAAAGTATCCAACATTGCAAGTAACCTTGTATCCCTCAAGCAATCCTGATGACATTGAT
CCTACACCTACTCTACTCTCTACTCTGATAAAAGCTATAATTCTGGAGTTAATATTCTACGCTGTTATTG
TCTGTGATTGGGTGTTGTAATTGTTAACTTTATTGTTAAGTACCACTTTGACCTTAACGAAGAAAAAAATCTTC
AAAGTAGACCTAGAAGAGAGTTTAAAACAAAACATGTAAGTAAAGGATAATTCTGAAATTAAAACACTCATGGATA
GTGTGATCATAAAATCTGATGTTGAAAGTGTGGGAGAGGACTTTGATTAAATAAAACACTCATGGATA
TGTAAGGAAACTGTCAAGATTAAATTGTTGATGTTGTTGAAAGTGTGAAAGAATAGTGTGACAAAG
ATCCTTTTCTACTGATACCTGGGTGTTGATATTATTGATGCAACAGTTCTGAAATGATATTCAAAATTGATCAA
GAATTAAAATCATCTATCTGACTGAGTCAAATACAAGTAAAGGAGGCCAAACAAACACATTGGGAAAAAA
AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNNTSFIKLNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTYLFE
ATEKRFKKNVSLIPENWKENPQYKRPKHENKHAADVIVAPPTLPGRDEPYTKQFTECGEKGHEY
IHFTPDLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCSAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFLLOQTENGWSWGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLLTDGEDNTASSCIDEVKQSGAIVHFAILGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISIWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNQDGVSRYFTAYTENGRYSLKVRAGGANTARLKLRRPPLNRAAYIPGWVVNGEIEANPP
RPEIDEDETTQTTLEDFSRTASGGAFVVSVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
FDVGKVQRYIIRISASILDRLDSFDDALQVNNTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPDKSHNSGVNISTLVLVIGSVVI
VNFILOTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCTCTAGGTGGAAACCCCTGGGAGTAGAGTACTGACAGCAAGACGGGAAAGACCATACTGCCCGGCAGGGTGA
CAACAGGTCTCATTTTGTACTCTCGTGTGGCTGCCTTCTATTCAAGGAAAGACGCCAAGGTAATTGGACCCA
GAGGAGCAATGATGTAGCACCTCTTCTTCTGGCTTAACCTCTTGTGGTGGAGGAGAACCTTGTGGGCTGCTCTTAGCA
ACTCAACACCAAGCGGCCCTCTCGGCTTAACCTCTTGTGGTGGAGGAGAACCTTGTGGGCTGCTCTTAGCA
GTGCTCAGAAGTGACTTGCTGAGGGTGGACCAGAAGAAAGGAAGGCTCCCTTGTGGCTGTGGCTGACATCAGAA
GGCTGTGATGGAAATGAAGGTAAAATCTGGAGATTCTACTCTAGTCATTGCTCTGCCCAAGAATCATCTTAAA
AGTAGAGAAGCTGCTGTGGTGTAACTCTTCAAGGAGGCAAGACTCTTCTAGAGGAAATGGATCAAGCACGTC
CGGGGGCCCCAAACGCGATGCTTCTTGTGGCTAGGCCAGGGAAAGCCCTTCCGTGGGGCCCCGGCTTGAGGGATGCC
ACCGGTTCTGGACGATGGCTGATTCCTGAATGATGATGATGATGTTGCTGGGGGGCTGCTTGCGTGGATTCCGGGGTGGT
GTTTGTGCTGTGCTCTCTGCTGTGATCTCTGTGCTGATCTGGCTGACATGTGGCTGCAACCCAAAAGGTGACGAGGAGCAC
CTGGCACTGCCACGGGCAACAGGGGAGGGGGTACCGGGCTTCTCAGGAGGAGCAC
CGCAACTACCTGAGCACCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCGAGGAGGAGTGACGAGCTCAGG
AATGGGCACTGATCAAGGCCAGCGATGCTGCTGGCTGGGCTCTGACAGGAGCCCCCAAGGAAACCCAGGCCAGCTC
CTGGGCTTCTGCACTCCAGGGCAACAGGAGGGTGAAGGTTGATCTGCGTCAAGCTGGCACAGAGTATGCGAGCTG
CTTCTGATAGCTTACTACAGAAGGTGATCAGCTGGAGATGCTGGGCTTACCCCAACCCGAGGAGAACGGCTGTG
AGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATTGAAATCAGCCTGGAGACCCCTGAACAATCTGAGAGAACAGC
CCCAATCACCGCTTACGGCTCTGATTTCTAGAAGGGATCTACCGAACAGAAAGGGCAACAGGACATTCT
GAGGTCACCTTCAAAAGGGACCAACAAACAGCAATTCAACCGCTCTTACATTCTTGGACCATTCAGGCCCCATCATGAAA
GTGAAAATGAAAGCTAACATGCCAACACGGCTTACATGTTATCTGCGCTCTGACAAAAGGGTGACAGTT
CGGCACTTATGCAAGAATTTCAGGGAGATGTGATTGAGCAGGATGGGAGAGTCCACTCTACTGTGTTACTTTGGG
AAAGAAGAAATAATGAAGTCAAAAGGAATCTGAAACACTTCAACAGGCTGCCAACCTCAGGAAACTTACCTC
CAGGCTGAATGGAAAATTCTCGGGAAGGGACTTGTGTTGGGGCGCTCTGGAGGGAAACAGCTCTTCT
TTTTCTGATGTTGGACATCTACTTCACATCTGAACTCTCTGACATACAGTGTAGGCTGAATACAGCCAGGGAAAG
GTATTTATCCAGTTCTTCAGTCAGTACAATCTGGCATAATATACGGCACCATGATGCACTCCCTCCCTGGAA
CAGGCACTGGTCTAAAGAAGGAAACTGGATTGGAGAGACATTGGATTTGGGAGTACGCTGTCGATATCGGTCAAC
TTCACTCATATAGGGTTGTGACATCAAAAGGTGGGGCGAGGAGATGTGACCTTATCGCAAGTATCT
CACAGCAACCTCATAGTGTACGGACCCCTGGCAGGACTCTTCCACCTCTGGCATGAGAACGCTGCACTGGACGAG
CTGACCCCCGAGCAGTACAAGATGTGCACTGAGCTTCAAGGCCATGAACGAGGCATCCACGGCACTGGCCTAGTGTG
GTGTTCAAGGACAGGATAGGGCTACCTTCAGGAAACAGGAAACAGAGAACAGTGAACAAAGCT
GAAGGATTGTGGGAGACATTTCCTTCTGCAATTACTGAAGAAGTGGCTGCAACAGAGAACAGACTCT
GGAGCACAAGAAATTGGACTGATGGTCAAGAGATGAGAACAGCTCGGATTCTCTGTTGGGTTTACACAGA
AATCAAATCTCCGTTGGCTGAAAGTAACCCAGTTGACCCCTGTGAGTGTCTGACAAAGGAGACAATGCTGTG
AGATTATAAGCTTAATGCTGGAGGTTTGATGTTGTTGAACTACAGAACACTGTTGTTTGTGTCATTTG
AATATTATCATGATTAAAGGAGCTTGTGTTGAAAATTCTAGGATCAAGGAGAACAGGAAACAGTGAATG
GCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAATATCAGAAGGGCAGGAGGAGATAGGCTTATTATGATACT
AGTGAGTACATTAAGTAAATAAAGGACAGAAGAAAGAACCTAAATATCTGTTGCTCATATTCCCCAAGAT
TAACCAAATAATCTGCTTATCTTGGTTGCTCTTAACTGTCTGGTTTCTTCTTATTTTATTAAGGAG
TTTTTCTCTGATGTTAGTCTGCTTATTAACTACAGCTTCAAGCCTTACAGAGAGCACAGTGGCCTAC
ATTTTTATTTTAAAGAAGATACTTGAGATGCAATTGAGAACCTTCTGAGTCAACAGCATCAAATTGATGCCATAT
CCAAGGAGCATGCCAAATGCTGTTCTGAGGACTGATGCTGAGCATTGAGACATGGGAAGGATGGTTGACT
AATACAGACGCTACAGATATTCTCTGAGAGTATTTCAGAAGGAGGCAACTGACAACTGGAGGGAAAGAAC
ACTTTCTGTTTACAGAAAAGGAAACTTACAGCTGGTATCTGATGTAACCTAAAGTCAGAACACCACATT
CTCCCTCAGAAGTAGGGACCGCTTCTACCTGTTAAATAACCAAAGTACCGTGTGACCCAAACATCTTTC
AAAACAGGGTCTCTGCTCTGCTTCTGCTTCTTACAGAAGGAAATATATATATATATATATATATATATG
GAAAGATCAATCTGCCAGAATCTAGTGGGATGAGGTTCTGACATGTTGATCTTACCCAGGGCCAGGGAG
TAACTGAAATTATTTTAAATAATGAGCTTACTCAATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CAAACATTTTAAATAATACAGTAAACATAGAGTGGTTCTCATTCTGTCAGGAAATATGAGGAGGAGGAG
ATTCAGTACGCTAAATTCTCTTGGCTTCTGCTTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG
CCATGAATGGAAGGGTGTATTGACAGCTAAATAAAATGATTTGTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRKQIAQLKEELQERSEQLRNGQQQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN
TLINVIVPLAKRVDKFRQFMQNREMCEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIOLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY
NPGIIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDDIKGWGGEDVHLYR
KYLHSNLIVVRTPVRLFHLWHEKRCMDELTPEQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGGCAGCAAGAGATTGTCCTGGGATCCA
GAAACCCATGATACCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCCCTGCTAGTCTCTAGTCTCAAATTCCAGTCCCCTGCACCCCTC
CTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTGATTGGATCCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATGGCCAGCCTTACCTGAGTGT
GGAAACAATGCCAGTGCCTCGATATTGACAGACAGTGTGACATTGACCCCTGATTTGCC
TGCTCTGCAGCCCCACGGATATGACCAAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCACCCCTGTATCTGGTGGACTTCCCGAAAATATGTAGCT
GCCAGCTCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGTCAGAACACCAGATCAACAG
TGAAGCCACATTGCAGAGCTCCACATTGTACATTGACTCTGATTCTATGACAGCTGAGTG
AGGCTGCTGAGAGGCCCTAGGGCCTGGCTGCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGTTATGAACACATTCTGAGTCATTGCATGAAAGTCAGGCATAAAGATCAGAACACCTC
AGTGCCTCCCTCAACCTAACAGAGCTGCTCCCCAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGCTACAACTCCCCCTGCTACCAGAGTGTGCTGGACAGTTTTATAGAAGGTCCCAG
ATTTCAATGGAACAGCTGAAAAGCTTCAGGGGACATTGTTCTCCACAGAACAGGAGCCCTCAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGGCATGGCTTGCTTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGAGGAATCTGGTTGGC
TGTCTCTGCCTCTCCTGGCTTTATTTCATTGCTAGAAAGATTGGAAGAAGAGGCTGGAAAA
CCGAAAGAGTGTGGCTTCACCTCAGCACAAGCCACGACTGAGGCATAAATTCTCTCAGATAC
CATGGATGTGGATGACTTCCCTCATGCCTATCAGGAAGCCTCTAAAGGGTGTAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCTCCCTGGACATCTCTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACCTGCAGAGCCTCAGCCTCTCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTGTTAATGCAGAGANAAACTCTGTTAGTTGCAGGGGAAGTTGGG
ATATACCCCAAAGTCCCTCACCCCTCACTTTATGCCCTTCCCTAGATATACTGCAGGGATCT
CTCCTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTGATCAATATATTGAAATTAAAG
TTCTGACTTT

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGCTCTGGCGCCCTGGGGACGGGCAGTTCCCTGTGTC
TCTGGTGGTTGCCTAAACCTGAAACATCACCTCTTATCCATCAACATGAAGAATGTCCCTACA
ATGGACTCCACCAAGAGGGCTTCAGGGAGTTAAAGTTACTTACACTGTGCAGTATTTCATCACAA
ATTGGCCCACCAAGAGGTGGACTGACTACAGATGAGAAGTCCATTCTGTTGTCCCTGACAGCTCC
AGAGAAGTGGAAAGAGAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCAGTGTGACCAACCAC
ACGCTGGTGCACCTGGCTGGAGCCGAACACTTTACTGCGTACACGTGGAGTCCTTCGTCC
AGGGCCCCCTGCCGTGCTCAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTCTGGTATGTTTGCCTATCTATTACCGTGTCTTT
TCTGTGATGGGTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATT
GATTGATTTATGAAATGAATTGACAAAGATTCTTGTGCCCTGCTGAAAAAATCGTGATTA
ACTTTATCACCTCAATATCTGGATGATTCTAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGTATCCAGCCTTAATGATCCTCAGCCCAGCAGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGGTATGCTCGATTGATGAAATTGACTCTGAAG
AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCAACTGACATTGCGGGGCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTATTGAGTCGCAGGCAGCGTGGCAG
TCCTGGCCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGCGACTGGGA
TCCCCAAACTGGCAGGCTGTATTCCCTCGCTGTCAGCTCAGGATTCAAGGGCTGCG
AGCCTCTGAGGGGATGGCTCGGAGAGGGTCTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCTAGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTCCTTGCCTTGTGCAAACAAAGTGAG
TCACCCCTTGATCCCAGCCATAAGTACCTGGGATGAAAGAAGTTTCCAGTTGTCACTG
CTGTGAGAATTACTTATTCTTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTA
GGTCTCTAACATGATGGTGGGCTCTGGAGTCCAGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCATAAAATGTTGCCAGACTGGGTGCAGAATTATTCAAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISSSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGYNEFDKRFVPAEK
IVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKGHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOEEVSTQGTLLSQQA
ALAVLGPQTLQYSYTPQLQDLDPLAQEHHTDSEEGPEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDSLGEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMCN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTCCCCAGGTCTGGCATCTGCAC TTGCTGCCCTTGACAC
CTGGGAAGATGGCCGGCCCCGTGGACCTTCACCCCTCTGTGGTTGCTGGCAGCCACCTTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGGAGCTGAAGGACCACAAAGCCACCAGCATCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTAACACCGCCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTGGCCAATGACCA
GGAGGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACCACTGAGGCCCAGCCACCATCCGCATGGACACCCAGTGCAAGTGGC
CCCACCCGCCCTGGTCCCTCAGTGA CTTGACCTGGGAGCCTGCGCATCCAAC TGCTGTA
TAAGCTCTCCTTCCTGGTGAACGCCCTAGCTAACGCAGGTCTGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTAAAAACCAAGCTGTGCTCCGTATCGAGGCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGTGACACCATTAGCTAACCTGAGCTGAGGCTGGCTGTGCTCTCCAGAAGA
AGGTGACCAAGTGGTTCAATAACTCTGAGCTGGTGAAGGCTGAGCTGGCTGTGCTCTCCAGAAGA
ATTCAATGGTCTGGTGGACTCTGTGCTCCGTAGAGTGCCCATGGCTGAAGTCAAGCATGGC
TGATCAATGAAAAGGCTGAGATAAGCTGGATCTACCCAGATCGTAAGATCTTAACACTCAGGAC
ACTCCGAGTTTTTATAGACCAAGGCCATTGCAAGGTGGCCCAACTGATCGTGTGGAGTGT
TCCCTCCAGTGAAGGCCCTCCGCCCTTGTGACCCCTGGGATCGAAGCCAGCTCGGAAGGCTCAGT
TTTACACCAAGGTGACCAACTTATACTCAACTTGAAATAACATCAGCTGTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGCCAAATTAGATCTGGGCTCCAGTGTGATTGGTGAAGG
CCTTGGGATTGGCTGGGAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCTCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAAA

FIGURE 78

MAGPWTFTLLCGLIAATLIQATLSPTAVLILGPVKIKEKLTQELDHNATSILQQPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPEAHRLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQILMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCTGAAGACGCTTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTCTGGGGCTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTGGTGCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACAGCACAGGCATACCCAGTGTGACATCTATAGC
ACCCCTCTGGGCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGTAGCAGGTGGAGTCATCCCTGGAGGCCCTGGGATTC
ATTCCCTGTTGCCATGGAATCTTACATGGGATCCTACGGGACTTCACTCACCACGGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGCTTTACTTGGCATTATTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTCTGCTCATCCCAGAGAAATCGCTCCAACACTACGATGCC
CAAGCCCAACCTCTGCCACAAGGAGCTCTCCAAGGCCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCCCTACAGCCTGACAGGGTATGTGTGAAGAACCAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCGAGGGCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTTCC
TCACCTTGCTGCCCTGCCCTAAGTCCCCAACCTCAACTTGAACCCCATCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCCTCTGGTTACCTGGACTCCATCCAAACCCACTAATCACA
TCCCACTGACTGACCCCTGTGATCAAAGACCCCTCTGGCTGAGGTTGGCTCTAGCTCATT
GCTGGGGATGGAGAGGAGAACAGTGGCTTTGTGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCATCCACTCTGTTATGACTCCACAGTGTCCA
GACTAATTGTGCATGAAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 80

MASLGLQLVGYILGLLGLLGLTVALPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITAQCD
IYSTLLGLPADIQAAQAMMVTSAASSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPIVAVNLHGILRDFYSPVPDSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCTCTCCCTCTGCTGGACCTCCTCGTCTCCATCTCTCCCTCTTC
 CCCCGTTCTCTTCCACCTTCTCTTCTTCCACCTAGACCTCCCTCCGTCCCTCCTTCC
 GCCCACCGCTGCTCCCTGGCCCTCTCCGACCCGCTAGCAGCAGACCTCCTGGGTCTGTGG
 GTTGATCTGTGGCCCTGTGCCTCCGTGCTCTCCCTCCGACTCCGCTCCGG
 ACCAGCGGCTGACCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCCCTCCCTGGTCTGGGA
 CTCGCGCTGCTCTGGTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
 CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
 ACTGCCTCGCCTGTACCTGCTCAGAGGGCGCCATGTGAGTTACCGCCTCCACTGTCCGCCT
 GTCCACTGCCCTCAGCAGCAATGCTGTCCAAGTGTGTGAAACCTCACAC
 TCCCTCTGGACTCCGGGCCACAAAGTCTGCCAGCACAAACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTTCCCTCCGCCGTGCCAACCAAGTGTGTCCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCCCTCACAAACCTGCCCGAACCAAGGTGCCAGCACCCCTCC
 ACTGCCAGACTCCTGCTGCAAGCCTGCAAAGATGAGGCAAGTGTGAGCAATCGGATGAAAGAGGACA
 GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGTGTCCTCTGCAGCTGC
 AGAGGCCGGCACCCAGCCCCACTGGCCTCAGGCCCTCTGAGCTTACCCCTGCCACT
 CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGCCTGAAGGAGAACATAAGAACGCT
 GTGTGCATGGCGGAAGACGTACTCCCACGGGAGGTGTGGCACCCGCCCTCCGTGCCTCGGC
 CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCGCCAGGACTGCCAGCGTGTGACCTGTCC
 CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTGCCAGAGG
 ACAAAAGCAGACCCCTGCCACAGTGTGAGATCAGTTCTACCAGGTGCTCAAGGCACCGGGCCGG
 CTCGTCCACACATCGGTATCCCCAAGCCAGACAACCTGCGTCGCTTGCCCTGGAACACGAGGC
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCAAGGCCACACAGCCAGAAATCTCCACTTGACTCAGATCAAGAAAGTCAG
 GAAGCAAGACTCCAGAAAGAGGCACAGCACTCCGACTGCTCGCTGGCCCCACGAAGGTCAC
 GGAACGTCTCCAGCCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTGAACCAAG
 ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
 TAAGAAGTTGCATTACCCCTAAAAAAAAAAAAAAAAAAAAA

23 25 27 29 30 32 34 36 38 39 40 42 44 46 48 50 52

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPAPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGT PAPTGLSAPIFIPRHFRPKGAGSTTVKIVLKEHKKACVHGGKTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTALPTARWPPRRSLERLPSPDPAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTGAGCATAGGCCAGTTCTCCTCTCTCTAATCCATCCGTACACCTCTCTGTCA
TCCGTTCCATGCCGTGAGGCCATTACAGAACACATCCTAGGCTCATGCTCAGTTGGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGCCAGACAAGGCTGTCCAGGCCCTGGTGGGGAG
GACGCAGCATTCTCCTGTTCCGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAACTGGTGAAGGATTCTATTGGGAGGGCCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGCCCTATGGGTCAGGATTAGTTCCAGTCTTACTACAGAAGGCCATCTGGAGCT
ACAGGTGTCAGCAGTGGCTCAGTTCTCATTTCATCAGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTTGAGATCTCTGACCGTCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTCGAGCCTATATCGGGCACCTGGCTACCAAAGTACTGGGAAACTCTGCTGTGGCCTA
TTTTGGCATTGTTGGACTGAAGATTTCTCTCAAATTCCAGTGAAATCCAGGCCAACTGGACTG
GACAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTACCCGAAGCTCGCCTCTGATCTGAAAAGTGTAAACCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCTGGGAGTGTGGCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTGTCTCCGATCATGGGACTGGGCTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCCGTTTATCAGCGTCTCCCCAGGACCCACCTACAAAAATAGGGTCTTCTGGACTA
TGAGTGTGGGACCATCTCTCTTCAACATAATGACCAGTCCCTTATTTACCCGACATGTCGGTTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGCTTATAATGAGCAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTTGCAAGGGCTCTGCAATCCAGAGACAAGCAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTCTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
CTCTCTTGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA
GATGAAGGGGACTGGCCTGTCACATGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGAGGGAAAGAAGG
CTGACATTACATTAGTTGCTCTCACTCCATCTGGCTAAGTGAATCTGAAATACACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTTCAATTACACTTCAGTAAAAAA

FIGURE 84

MAIMLSLVSLKLGSQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRRFRGQFSSVVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVP LISITGYVDRD IQLLCQSSGWFPRTAKWKGPGQDLSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAE LDWRRRKHGQAE LR D A R K H A V E V T L D P E T A H P K L C V S D L K T V T H R K A P Q E V P H S E K R F
TRKS VVA S Q S F Q A G K H Y W E V D G G H N K R W R V G V C R D D V D R R K E Y V T L S P D H G Y W V L R L N G E H L Y F T
LN P R F I S V F P R T P P T K I G V F L D Y E C G T I S F N I N D Q S L I Y T L T C R F E G L L R P Y I E Y P S Y N E Q N G T
P I V I C P V T Q E S E K E A S W Q R A S A I P E T S N S E S S S Q A T T P F L P R G E M

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTCCGTGA
CGGTGAGGAAGGCTGTGTCCATGTGCCTGCTCCTCTCCTACCCTCGCATGGCTGGATT
TACCCTGGCCCAGTGTTCATGGCTGGGCCAGGGGGCAATACAGACCAGGATGCTCC
AGTGGCCCACAAACAACCCCAGCTGGGCAGTGTGGGAGGACTCGGACCGATTCCACCTGTG
GGGACCCCACATACCAAGAATuTGCACCCTGAGCATCAGAAGTGCAGAAGTGATGCGGGGAGA
TACTTTTCGTTGGGAGAAGGAAGTTAAAATGGAATTTAAAATCACCCGCTCTGTGAA
TGTGACAGCTTGACCCCACAGGGCCAACATCCTCATCCAGGACCCCTGGAGTCCGGTGCCCC
AGAATTGACCTGTCTGTGCCCTGGCTTGTGAGGAGGGACCCCTTATGATTCTCTGGATA
GGGACCTCCGTTGTCCCCCCCTGGACCCCTCCACCCACCCGCTCCCGGTGCTACCCTCACCCA
GCCCCAGGACCCATGGCACCAGGCCTCACCTGTCAGGTGACCTCCCCCTGGGGCAGGTGACCAGA
ACAAGACCGTCCCATTCAACGTTCCTACCCGCCTCAGAACTTGACATGACTTGTCCAAGG
GACGGCACAGTATCCACAGTCTGGAAATGGCTACTCTGTCACTCCAGGGCCAGTCTCT
GCGCTGGTCTGTGCAGTTGATGCAGTGACAGCAATCCCCTGGCAGGCTGAGCTGGA
GAGGGCCTGACCCTGTGCCCCCTCAGCCCTCAAAACCCGGGGTGCTGGAGGCTGGCAC
CTGAGGGGATGCAGTGATTCACCTGCAGGCTCAGACCCTCTCGGTCTCAGCAGGTCTACCT
GAACGTTCCCCTGTCAGGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTGGAGGCTGGGA
CCACAGCCCTGGTCTCCCTGTCTTTGTCGTCATCTCGTGTAGTGAGGTCCTGCAGGAAGAA
TCGGCAAGGCCAGCAGCGGGGTGGAGATCGGGCATAGGGGATGCAAAGGCTGTCAGGGGTTC
AGCCCTCTCAGGGGCCCTGTACTGAACCTGGCAGAAGGACAGTCCCCCCAGACCCAGCAGCCTCCCC
CTTCGTCCCGCTCTCAGTGGGGGAAGGGGAGGCTCCAGTATGCATCCCTCAGCTCCAGATGGT
AAGCCTTGGGACTCGCGGGGACAGGGGCCACTGACACCCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAACTGCAGAGACTACCCCTGATTGAGGGAGAGTC
GAGGCTGATTTGAGATTACAGCCTCAACGTGATGAGCTAGTATGATTGCC
TCCCCTTTTTTTTTTTTAATAAAAAGACAGACAAAATCCTA

FIGURE 86

MLLLLLPLLWGRERAEQGTSKLLTMQSSVTQEGLCVHPCFSYPSHGWYIPGPVVGWYFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRDAARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSVWPACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFFGASVTTNKTVHLNVSYPPQNLMTVFQGDGTSTVLGNGL
SLPEGQSLRLVCAVDADVSNNPAPRLSLSWRGLTLCPSQPSNPVGULELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLSQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGSSASQGPLTEPWAEDSPPDQPPPASARSSVGEQELQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGTAC
CCAAGGAAAGTCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTCCTGCTGTTTC
TCATAGCGACCACCAGAGGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA
TGGCCTGTATTTCTCCGACTGAGAATGGTGTATCTACCAAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGGC
CAACTACAACACCTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCAGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCCCTAT
GGCCAGCGGGAAATTCACTGCGGGATTGTTCAAGTGCAGGGTATTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAACCCAGACCTCTCCCAACCAGATGAGATCCAAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQILSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIKVYDFGDAQKTASYYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGTTGCGGGGAGACTTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTGTCCCCGAGGGCATGGCCGGGTCTCAGGGTTGTGCCCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCCCGGACAGCAACATAAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACCCCTGGGCCTCTTGCAGAGTGGAGCTGGCCGTTCCCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCCTTCAT
ATTGAGCGTTGGAGTGCACTACGTATTGGTACATTTTGTCAGTGCAGTGCCTTCCAGCTG
TCACTGAAATGGCTTATTGTCACCGCTTGGCTGAAAAAGAAACCTCTGATTACCTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCGCTCGTATTCTGGAAAGAAGGAAG
GCATAGGCTCGGTTTCCCTCGAAACTGCTCTGCTGGAGGATATGTGTTGAATAAAATATGTTGTAGTAACA
TCTTGAGTCTGGATTATCCGATTGTATTAGTGCTTGTAATAAAATATGTTGTAGTAACA
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCAGAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGAATGCAACTGAC
 TCGCTGCTGCTCGTGTTCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
 GTCTCCCGGCTCAGAGGACCTGAGCGTATGACCACGAGGGCAGCCCCGGCCGGTGCCT
 CGGAAGCGGGGCCACATCTCACCTAACGTTCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
 GGCCCCGCCTGGGAGGCTGGGCATTCTGGCAGCCCCCAACGCCGAACCACAGCCCC
 CACCCCTCAGCCAAGGTGAAGAAAATTTGGCTGGCGACTTCAACTTCAACATCAAGACGGT
 GCGCTGAACCTGCTCGTCAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
 CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCTCCAGTAAAGCTGTAG
 AGTTCACCCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAATCTCAACTGCCGGATG
 GAGTGGGAGAAGGTAGAACGGGCCGGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
 CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTCAAAGTCGTCTGTG
 TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAAC
 CATAGTGATACCCCTACTACCCATCTGGGTGACCCGGGCAGGCCACAGAGGCCAGGCCAGGGC
 TGGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCAGGGCAGGGAAAGGGTTGGCCTC
 AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCAAGTCTCAAGTGGCAG
 AGAAAGGGTCCAAGTGTGGTCCAACCTGAAGCTGTGGAGTGAAGTACAGGAGCACTGG
 AGGAGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGG
 TCCCGAGGCCCTGTGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAC
 CCTTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATTTCATCAGTGTGGACA
 GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTCTAGGAGCCAGTCAGCAGGGTGGGG
 GCCAGAGGAGCTCCAGCCTGCCAGTGGCGCCCTGAGCCCTTGCGTGTGAGCATGG
 CATGAGGCTGAAGTGGCAACCCCTGGGTCTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
 CAGGCCACCCCTTCCAAAATTCCCTCTGCCAGTACTCCCCCTGTACCCACATTGCTGATG
 GCACACCCATCTTAAGCTAACAGACAGGACATTGTGGCTCTCCACACTAACGCCACAGCCC
 CGCGTGTGTGTGTCCTCTCCACCCCAACCCCTGCTGGCTCTCTGGAGCATCCATGTCCCG
 GAGAGGGTCCCTAACAGTCAGCCTCACCTGTCAAGACCGGGGTTCTCCGGATCTGGATGGCG
 CGCCCTCTCAGCAGCGGGCACGGTGGGGCGGGGCCAGAGCATGTGCTGGATCTGTC
 TGTGTGTCTGTGTGGTGGGGAGGGAGGGAGTCTTGAAACCGCTGATTGCTGACTTT
 TGTGTGAAGAATCGTGTCTTGAGCAGGAATAAGCTGCCCGGGCA

33 433 53 63 733 83 933 103 113 123 133 143 153 163 173

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVANLLVTGKIVDHGNGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPKVVCVYIAFYSTDYRLVQKVCVDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGC~~GGCGT~~GTTCTTCGGCTGCCCTCATTGCC~~TT~~CGGGCTGCCGCTGCC
CTTATGTCTCACCATGCCATGAGCCGTTGCGTATCATCTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATT~~CGT~~CCCTGTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGATC~~TTGGAGCGTTGTCTGT~~TATATCCAAGAA
ATGTTCCGATTGCATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGCC~~T~~ATGTTCTGGCTGGCTTGGATCATGA
GTGGAGTATTCCTTGTGAATACCCATCTGACTCCTGGGCCAGGCACAGTGGCATT~~CAT~~
GGAGATTCTCC~~T~~CAATTCTCCTTATT~~CAG~~CTTCATGACGCTGGT~~CATT~~ATCTGCTGCATGT
ATTCTGGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGT~~CAG~~CCCAGAC~~CT~~CATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGCTCATGGCAC~~TGGCATT~~CTAGCTGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCC~~T~~GCTGCCAAGACAAGAACTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGACCAGC~~ACT~~CCAAACCGCAGACTACATCTTAGAGGAAGCACA~~ACT~~TGTGCCT
TTTCTGAAAATCCCTTTCTGGTGGATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLIAYVSGLGFMSGV
FSFVNTLSDSLGP GTVGIHGDS PQFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKW GILLIVLLT
HLLVSAQTFI SYYGINLASAFIILVLMGTWAFLAAGGSCRS LKLC QDKNFL LYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

FIGURE 96

MRSTIILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTPAGIQRSTHAIIEATTESANGIQ

Signal peptide:
amino acids 1-16

FIGURE 97

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLPEPIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRAGQNILLVILAMDLAGAEGGSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRIMDFAIERGDTEGTFLDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGPAGATATVTVLVERVMPPPQLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCTIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTLPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTLALHWVEP
REHIIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKMPTKLSAVGILVGTLVAIGIFLI
LIFTHTWMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCCTGGAGGAAGCTAAGGAACCCAGGCATCCAGCTGCCACGCCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCTGGAAAGCACCAGCCTTTA
TCTCTTCACTTCAGTCCCCCTTCTCAAGAATCCTCTGTTCTTGCCCTCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATTTGCTTCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTATGTTGGTCTACTATTGCAATTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGATCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGTCA
GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTTGAG
TTCAGCACAGCTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCCAGTGGGCCAGCACAGTCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACCTCCAGTGGGCTAGCAGCACAGCCACCAACTCTGAGTCCAGCACA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCACCTCCAGTGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTAGGGCCAACACAGCCACCAACTCTGAGTCCAGTGGGCCAGCACTGCCACCAACTCTGAGTCCAGC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCCAGTGGGCCAGCACAGCCAC
GCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAAACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGTCCAGTGTGACCTCTGAGGCTCTGGAACAGCAGCAGCTGACTGG
ATGACACACAACCTCCCATAGTCATCTACTGAGTGTGAGGCAAAGCCTGGTGGTCCCTGGT
GCCGTGGAAATCTTCCTCATCACCCTGGTCTCGGTTGTGGGCCGTGGGCTCTTGCTGGC
TCTTCTTCTGTGAGAAACAGCCTGTCCTGAGAAACACCTTTAACACAGCTGTCTACCACCC
CATGGCCTCAACCATGGCCTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCAGAGGCCAG
GTGGAGTCTTAAGTGGTTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGGCCAG
ACAGGGGGCCCTGAGCAGCCCCGGAAAGCAAGTCCCCATTCTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCTTCATTCATCCAGGAGACCCCTCCAGCTTGTGAGATCCTGAA
AACTTGAGAAGGTTATTCTCACCTTCTGCTTACACAGACACTGGAAAGAGAAACTATAT
TGCTCATTAGCTAAGAAATAATACATCTCATCTAACACACAGCACAAAGAGAAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTGAGATGAACCTCAGTTAGGAGAAACCTCATGCTGGACTC
CATCTGGCATTCAAATCTCACAGTAAAATCAAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTSSGVSTATNSESSTTSSGASTATNSDSSTTSEASSTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNSLRLNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVPSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCCTCGCGTTACGGGATGAATTAACGGCGGGTTCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTGCCACGCACCCACTCGCGTCGCGCGCGTGCCTGCTGTACAGGTG
GGAGGGCTGAAACTATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGAAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGAAAGGGAAATTGTGGTGCCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTCATAAAGCAGGACTCTAAAGCTTGAAATCAT
GGTGTCAAGGAAAGGGATTACTTACTGACTCTGTTGGGAAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTGTGGCAACATGGCTCACCCACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGAAATTGCGTGTGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAGGTGTTCTGGATTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCAAGGAAATGAAAGGATGACAAGAGCCTTCGAAGACATGATTGATTACTTTGT
ATATTCAACGAACCACTTCAACTCCTCATATTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGTTTACTTTGTGGTAGACCGTCAAGAGAAGGTAAGAACCTGATGCTGTCCATG
ATATCACTGTGGCTATCCTCACAAACATTCTCAATCAGAGAACGACCTCCAGGGACTTT
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGAAAGAGAAAGAAGAGGCTGCGTCCCTATCAAG
GGGAGAAGAATTTTATTACCGGACAGAGTGTCACTCCACCTTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCTGTTCAAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTAAAGTGGTATTTATAATCACCATTGAATCTTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACATTGCACTGATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTAAAGGTTGCATGTGAAAACCTAGAGCATATTTG
GAAATGTTCTAAACCTTCTAAGCTCAGATGCATTGCACTGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTCGACTTAATTGTGGAAAAATATTGCTACAATT
TTTAATCTCTGAATGTAATTGACTGTGTACATAGCAGGGAGTGATGGGGTGAATAACTT
GGGCCAGAATATTAAACAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIFYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKLASKLGVPFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRSNAFAEKNGLQKYEYVLHPRTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEI1ELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAATCATCCATCCACCCCTGCTGTCATCTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTTGGTCTCAGTTCTACGAGCTGGTCAAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCAGGCCTGGTGGGGAGGACGCCGTGTTCTCTGCTCCCTCT
TTCTGAGACCACTGAGAGCTATGGAAGTGCAGGGTCTCAGGAATCAGTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAACGAGATGCCACAGTATCGAGGGAGAACTGA
GTTCTGGAAGGACTCCATTGCGAGGGGGCGTGTCTCTAACGGTAAAAAACATCACTCCCTCG
ACATCGGCTGTATGGGTGCTGGTCAAGGTTACGATGAGGAGGGCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTACTTCTCTCATTTCATCGTGGATATGTTGACGGAGGTATCCA
GTTACTCTGCTGTCTCAGGCTGGTCCCCAGCCCACAGCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTCAGACTCCAGAGCAATGCAGATGGTACAGGCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAATGCTGGGAGCATATTGTTCCATCCACCTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTAATTGATAGGAGAGACGTTTCCAGCCCTCACCTGGCCTGGCTTCTA
TTTACTGGTTACTCTGTTGCCCCGTGTGGTGTGGTATGGGATGATAATTGTTCTTC
AAATCCAAGGGAAATCCAGCGGAACCTGGACTGGAGAAGAACAGCAGGACAGGAGAAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTACTGGATCCAGAGACGGGACAGCTCTGCG
TTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCAGGGAGGTGCTCACTGAGAAGAGA
TTTACAAGGAAGAGTGTGGCTCTCAGGGTTTCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGAGGGTGTGGTATGGGAGTGTGGTGGGATGACGTAGACAGGGGGAGAAC
ATGTGACTTTGTCCTCCAACAATGGTATTGGGTCTCAGACTGACAACAGAACATTTGATTTC
ACATTCAATCCCCATTTATCAGCCTCCCCCCCAGCACCCCTCTACACGAGTAGGGGTCTCCT
GGACTATGAGGGTGGGACCATCTCTCTTCAATACAATGACCAAGTCCCTTATTATACCCCTGC
TGACATGTCAGTTGAAGGCTTGTGAGCCATATCCAGCATGCGATGATGACGAGGAAAG
GGGACTCCCATATTCAATGTCAGTGTGGGGATGAGACAGAGAACGACAGGCTTAAAGGGC
CCACACCAACAGACCCAGACACGCCAGGGAGAGTGTCCCAGGGTGGCCAGCTTCTCT
CCGGAGCCTGCCACAGAGAGTCAGCCCCCACTCTCTTGGGAGCTGAGGTTCTCTGCC
TGAGCCCTGCAGCAGCGGAGTCACAGCTCCAGATGAGGGGGATTGGCTGACCTGTGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGACGGAATAGACTCACATTAGTTAGTTGTAAAA
CTCCATCCAGCTAAGCGATCTGAACAAGTCACAACCTCCAGGCTCCTCATTTGCTAGTCACGG
ACAGTGAATTCTGCCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGAGGTT
TGAGGGCAGGTGTTGCTAATGATGTTTTATATTACATTCCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTATACCAAATCACCAGGGATAGTTATTGAACACC
TGCTTGTGAGGCTAAAGAATAAAGAGGAGGTAGGATTTCAGTGAATTCTATAAGCCAGCAT
TACCTGATACCAAAACAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACACTACAGGTCCATATCC
CTCATTAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATTTA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGAGGTTAATATTAAAT
ATCAACCAGTGTAAATTCAAGCACATTAATAAGAAAAAGAAAACCATAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKVFQALVGEDAVFSCSLFPETSAEAMEVRRFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCSSGWFPQPTAKWKGPGQGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHIAEQSHEVESKVLIGETFFQPSPWRLASIILGLLCGALCGVVMGMIIVFFK
SKGKIQAEILDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWWVGCRDDVDRGKNNVTLSPNNGYWVRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGGAACCCGGGTTATCGGCCTCGTCATCTCATATCCCTGATTGTCCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAACACTACAATTACTATAGCACATTGTCATTACAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTCTAACAAATTTCAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAGGAAATGCATTATAAAATCTCCATTAAAGGGAAGAATTGTCAGTCTCAGGTTATCAAGTTC
AGTCACAGAACATGGAGTGTGGCTCATATGCTGTTGATTTGTAGATTTCACTACTGAGGATCCTGA
AACTGTAGATAAAATTGTCACCTGTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAATTAAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
ACACGAAGAAGTAAACTCTAGGTCAAGTCTCAGGATCGTGGGAGCAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAGCCACATGGC
TTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCCCTGCCAGATGGACTGCTTCCTTGGAGTAACA
ATAAAACCTCGAAATGAAACGGGTCTCCGGAGAATAATTGTCATGAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGCAGAGCTTCTAGCCCTGTTCCCTACAAATGCACTAGATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGACAGGATTGGAGCACTGAAAAAT
GATGGGTACAGTCAAAATCATCTCGACAAGCACAGGTGACTCTCATAGCGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTCAGGCTTAGAAGGAAAACAGATGCAT
GCCAGGGTACTCTGGAGGACCACTGGTAGTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
AGCTGGGAGATGAATGCGAACCCAAACAAGCCTGGTGTATTACTAGAGTTACGGCCTGGGGACTG
GATTACTCAAAATGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTGTTGTTGACATAATTGTAAT
GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGTTGATGCACTGATGTTCTTCTCCAGCTCTGTTCCGACGTAAAGCATCCTGCTTGTCCA
GATCAACTCTGTCATCTGAGCAATAGTTGAAACTTATGTCAGAATTGTTGACATAAGATAATACAATTAC
ATTACAGCCTGTATTCTGTTCTAGAAGTTGTCAGAATTGTTGACATAATTGTAAT
GCATATATACAATTGAAAGCACTCCTTCTTCAGTTCTCAGCTCCTCTCATTTCAGCAAATATCCATT
TCAAGGTGCAAGAACAGAGTGAAGAAAATATAAGAAGAAAAAAATCCCTACATTGTTATTGGCACAGAA
AAGTATTAGGTGTTCTTAGGAAATTAGAAATGATCATATTCAATTGAAAGGTCAAGCAAAGACA
GCAGAAATACCAATCACTCATTTAGGAACATGGGAACTAAGTTAGGAAGTCCAGAAAAGAGCCAAG
ATATATCCTTATTTCATTCCAAACAACACTATGATAAAATGTGAAGAAGATTCTGTTGACCT
ATAATAATTACAAACTTCATGCAATGACTGTTCTAAGCAAATTAAAGCAAATATTATTAACATTG
TTACTGAGGATGTCACATATAACAATAAAATATAACACCA

FIGURE 106

MMYRPDVVRARKRVCWEPWIVGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPREEFVKSQVIKFSSQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATLVSAAHCFTTYKNPARWTASFGVTIKPSKMKGRL
RRIIVHEKYKHPSHDYDISLAELESSPVPTNAHRVCLPDASYEFGPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWIYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAAGAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCGCGAAGAAGTCCCTG
CCCCGATGAGCCCCCGCCGTGCGTCCCGACTATCCCCAGGCGGCGTGGGGCACCGGGCCAGC
GCCGACGATCGTGCCTGGAGTAGGGATGTGGTAAAGGATGGGGCTCTCCCTT
ACGGGGCTCACAAATGGCCAGAGAAGATCCGTGAAGTGTCTCGCCTGCCTGCTACGCCCTCAA
TCTGCTCTTTGGTTAATGTCCATCAGTGTGGCAGTTCTGCTTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTCCT
GTGTTCATCCGGTCACTGTTGCTTGTCTTCTTATCATTGTTGGGGATGTTAGGATATTG
TGGAACGGTGAAGAAGAAATCTGTGCTCTTGATGGTACTGGGAAGTTGCTTGTCACTTCT
GTGAGAACTGGCTGTGGCTTGGACATAGAACAGGAACCTATGGTTCCAGTACAATGGTC
GATATGGTCACTTGAAAGCCAGGATGACAATTATGGATTACTAGATATCGGTGGCTACTCA
TGCTTGGAAATTTTTCAGAGAGAGTTAAGTGCTGGAGTAGTATATTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTTAGAGAATTCCAGGATGTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTTGTTGGAGAAAATGTATTCCCT
TTTGAGAGGAACCAAACAAC~~T~~GCAGGTGCTGAGGTTCTGGGAATCTCATGGGGTACACAAA
TCCGGCCATGATTCTCACCATTACTCTGCTCTGGCTCTGTATTATGATAGAACGGGAGCCTGGG
ACAGACCAAATGATGTCCTGAAAGATGACAAC~~T~~TCAGCACCTGTATGTCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAGAATCTTGAACACACATCCATGGCAAAACAGCTTAATACACACT
TTGAGATGGAGGAGTTAAAAAGAAATGTACAGAAGAAAACACAAACTGTTTATTGGACT
TGTGAATTGGAGTACATACTATGTGTTCAGAAATATGAGAAATAAAATGTTGCCATAAAA
TAACACCTAACGATATACTATTCTATGCTTAAATGAGGATGGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTAAATGCTGAAGACAGATGTCAACCCACTGTGTAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTGAGGCACTGGTTGATTAGCATTTCCGCA
TCCATGCAAACGAGTCACATATGGTGGACTGGAGGCATAGTAAAGGTTGATTTACTCTACCAA
CTAGTATATAAGACTAATTAAGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
CTCAGCGATCTTCTGATGCTAAATAAAATTATCTACAGAAAACTTCAATATTGGTGA
ACCTAAATGTGATTGGTCTGGTTACTAAATATTCTACCACTAAAGAGAACGTAACACAT
TGTCTTAAGCTGATCAGGGATTGGTATATAAGTCTGTGTTAAATCTGTATAATTCACTCGAT
TTCAGTTCTGATAATGTTAGAATAACCATGAAAGAAAATTGTCCTGTATAGCATCATT
ATTGGTCTTCCTGTTAATAAGCTTACTATTCTGCTCTGGCTTATATTACACATATAAC
TGTGTTAAACTAACACTAATTGAAAGGAAATTACAGTGTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGCTTTAGGAAGTATAAGAAAATTGACACATAACTTAGTTGATTCA
AAGGACTTGATGCTGTTCTCCAAATGAAGACTCTTTGACACTAAACACTTTTAAAGAAA
GCTTATCTTGCCCTCTCCAAACAAGAAGCAATAGTCTCCAAGTCATATAAAATTCTACAGAAAA
TAGTGTCTTCTCCAGAAAATGCTGTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTGTGTTATTCACTGATTAATATGTTGAGGAAATTACACAGATTATAAAATTGTTTACAA
GAGTATAGTATATTGAAATGGGAAAGTGCATTACTGTGTTGTTGATTTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTGTCAATAATATTCTAGAGAGTAA

.....

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTATRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLNLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACCTCATCCTCATCCTCTTCCCTGTATAAAGCCCTACCAGTGCT
GATAAAAGTCTTCTCGTAGAGGCCCTAGAGGCCCTAAAAAAAGTCTGAAAGAGAAGGGACAAGGAACA
CCAGTATTAAAGAGGATTTCCAGTGTGTTCTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTCACCTG
CCTCTTCATCACAGGCACCTCCGTACCCGTGGCCTAGATCCTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCACTGGATGAGTCTCAAGGTCCCTATGTGACAACCATGTGAATGGGAG
TGGTACCACTTCACGGGCATGGGGAGATGCCATGCCATCCTCTGCATACCAAGAAAACACTGTGGAACCCA
CGCACCTGCTGGCTCATGGCAGGCCACCCCTAGAAGGCCAGGGATTGTGCAACGCCAGGCTTGCCAGCT
TCAATGGGAACTGCTGTCTGGAACACCACGGTGGAAAGTCAGGCTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTCCACGTACTGTGGTCACTTTATGACATCTGCAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTCACATGCCCTCAGGAACCTGTCTAGGCCCTGACAGGCAGACATGCTTG
ATGAAAATGAATGTGAGCAAACAAACGGTGGCTGCAGTGAGATCTGTGTAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCGTGTGCTAAGAAGTGTGAGGACTTGTGAAGACGTTGAAGGATGCCACATAACAA
TGGTGGCTGCAGCCACTTGCCTGGATCTGAGAAAGGCTACCACTGTGAATGTCCCCGGGCCTGGTGT
CTGAGGATAACCAACACTTGCCTGGTGTGCAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTCTCCCTGACCAACACCTCCGCCAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGTACACTGGTCATGTGGTGAATGACAAGATTGTGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCGGGAGCAGCGGGACTTCATCATCCGAACCGACAAGCTGCTGATCCGGTG
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACGTTCCAACTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCCAATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCTCAAGCTTGTGACTCCCTACTTGGCATTGAGCCGTGGTGCACCTGAGCGGC
TTGAAAGCTTGGAGAGCTGTTGCCACCCCACTTCAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGTAAAGCAGTACACATCCGGGATCACCTAGCAAAGCACTCCAGG
TCCCTGTCTCAAGTTGTGGCAAGACCAAGGAATGTTCTGCACTGCCGGTCTTGTCTGTGGAGTG
TTGGACGAGCTTCCCGTGTGCCAGGGTGCACCGCGAATCGTCGTGGGCAGGAGGAGGACTCAGC
CGGTCTACAGGGCAGACGCTAACAGCGCCGATCCGCATCGACTGGGAGACTAGTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTGCTTGGAGCTTCTCCCTACCCCTCAAGAACATCTGCCAACAGC
TGGGTCAGACTTCACACTGTGAGTTCAAGACTCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTACAGCACTGCTGAACATGTGGCTGGGTGGGTTTCATCTTCTAGGGTTAAAAACTAAACTGTCCA
CCCAGAAAGACACTCACCCCTTCCCTACACTTAAATACCTCGTATGGTGCACAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCTAGAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCCAAATGACTCCTAAATATGTAATTATGTTACCTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAATTGGAAAGTGTATCAATAAACAGTATATAATT

FIGURE 110

MPPFLLLTCFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPEHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGYYYVRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSHSCLGSEKGYQCECPRLVLSEDNHCTQVPVLCKSNAIEVNIPRELVGG
LELFLLNTSCRGVSNNGTHVNILFSLKTCGTVVDDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPV
CEFFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYLYLIRDGVSVDDSVKQYTSRDHLAKHFQVPVFKFVGKDHEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGTGACCAAGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGAGGCCCT
CCTGCTGCCCTGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGC**ATGT**
TACAGGATCCTGACAGTGTGACACCTCTGAACAGCCTCGATGTCAAAACCCCTGCGAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAAACTACTTCCTCTGGGGCAGCCTCTCACTTCATCCGA
GGAAGCAGCTGTGACGGAGAGCTGGACTGTCCTGGGGAGGACCGAGGACTGTGCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCAAGGACCGATCCACACTGCAGGTGCTGGACTGGCCAC
AGGGAACTGGTCTCTGCCTGTTCGACAACCTCACAGAAGCTCTCCCTGAGACAGCCTGAGGAGATGG
GCTCACAGCAGAGCTGTGGAGATTGGCCAGACCAAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGATGCGGAACATCAAGTGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTTGCCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGGAGGAGGACCGCTCTGTGGATTCTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGAGCATCCTGGACCCCCACTGGTCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGGGGCAGGCTCAGACAAACTGGCAG
CTTCCCACCCCTGGCTGTGCCAAGATCATCATCATTGAATTCAACCCATGTACCCAAAGACAAATGACA
TCGCCCTCATGAAGCTGCAGTCCCACTCACCTTCAGGCACAGTCAGGCCATCTGTCTGCCCTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGCTGACATACTGCTGCAGGCGTCAGTCCAGGTATTGACAGCACAGGTGCAATGCAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTTGAGGAGCATCCGGAAAGGGGTGTTGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTTACCAATCTGACCAAGTGGCAGTGTGGTGGGATCGTTAGCTGGGCTATGG
CTGGGGGGGCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCACTGGATCTACAATGTCT
GGAAGGCTGAGCTG**TAAT**GCTGCCCTTGCAGTGTGGAGGCCCTCCTGCCCTGCCACCT
GGGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCCGGTACACCCCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGGCTCAATTCTGTAAGAGACCCCTGCAGCCAGGGGCCAGAGGAAGTCA
GCAGCCCTAGCTGGCCACACTTGGTGTCCCAGCATCCAGGGAGAGACACAGCCCACGTGAACAAAGGTCT
CAGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGCTGCGCCAGCCCTGCGCTTCACCCATCCCCAA
GCCTACTAGAGCAAGAAACCGAGTGTAAATATAAAATGCACTGCCCTACTGTGGTATGACTACCGTTACCT
ACTGTTGTCTTATTACAGCTATGCCACTATTATAAGAGCTGTGTAACATCTCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLKVILDKYFLCG
QPLHFIPRKQLCDGEELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAEATACRQMGYSRAVEIGPDQDLDVVEITENSQELRMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVEVGEEASVDSWPWQVSIQYDKQHVCAGSILDPHWVLTAAHCFRKHTDVFNWKVRAAGSDKL
GSFPSLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPPFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTQCQGDGGPLMYQS
DQWHVVGIVSWGYZGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTCAGCAACTAAAAAGCCACAGGAGT
TGAACACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCTACTCCTACCTACATTAAAC
TGTTTTTGTCTTGTAACTAGCCTTACCTTCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGAACGAGAACAGAGGTTCTACCCACACCGTCCCCTCGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCTGCTGGAGCAGTGCCCTCACCAACTGTCTACGTCT
GGAGGCAGTGACTCGGGCAGTGCAGGTAGCTGAGCCTTGGTAGCTCGGGCTTCAAGGTGGC
CTTGCCCTGGCGTAGAAGGGATTGACAAGCCGAAGATTCATAGGCATGGCTCCACTGCC
AGGCATCAGCCTGCTGTAGTCATCACTGCCTGGGCCAGGACGGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTTATGTCTGCACATCACCTG
ATCCATGGCTAACTGAACCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAAACATGAGG
CAGAAGGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAAGGAACTTGTGCCAAATTATGGGTCAAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGCATTCAGTGGACATGTGGGGAAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGCTGGAAAGGAGCAAGCAAAGTGACCATTCTCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTCAGCAGCAAGGCCTGAGAGCTGATCAGAAGGCCTGCT
GTGCGAACACGAAATGCCCTCAGTAAGCACAGGCTGAAAATCCCCAGGAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGCCCCAAGACCAAAAGGAGCTAGAGCTGGTT
CAAATGATCTCCAAGGGCCCTTATACCCCAAGGAGACTTGATTGAATTGAAACCCAAATCCA
AACCTAAGAACAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTTGGCTGTAATGCCAACAT
TTGGGAGGCCAGGCGGGTAGATCACCTGAGGTCAGGAGTCAAGACACCAGCCTGGCCAACATGG
TGAAACCCCTGCTCTACTAAAAATACAAAAAAACTAGCCAGGCATGGTGGTGTGCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTGAACTGGAGGTGAAGGAGGCTGAGACA
GGAGAACACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAGAATTA
TGGTTATTGTAA

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLTEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTAGCCTCTCAAAGCAAGGAAAGACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCCT
GTTTGCCCCAGCAAGCACTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAGTGACGACTTTAAAACGGATAACTGG
CATCTACTTCGTGGTCTTCAAAATGTTTATCAGATTAAAGTGATTCCCTGAATT
CTGAACCCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAACAGTCAGTG
ATTTGGGTCCCGAGCAAGAAAGCCTATTGAAAACCGAGATTCTTAAATCCTGA
GATTGTGATAACGTGACCATGTATTGATCAATCCCCTCTAATATCAGTTCTGAGTTACAAG
ACTTGAGGAGGGAGGGAGAAGATCTTCACTTCCCTGCCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTTCAAATAATGACTATACTGAAAATGGAATAGAATTGATCCCCTGCTGGATGAGAGAG
GTTATTGTTGTATTACTGCCGTGAGGCAACCGCTATTGCCGCCGCTGTGAACCTTACTA
GGCTACTACCCATATCCACTGCTACCAAGGAGGACGAGTCATCTGTCGTCTCATGCCCTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTAAAGGAGGTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCAGAATTACTTGTAGGTAATTCTCTCTCATGTTCTAATAAAACTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEETTTFFEQSVIWVPAEKPIENRDFLKN SKILEICDNVTMYWINPTLISVSE
LQDFEEEEDILHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCCTCAGGGAGCGCGTTAGCTTCACACCTCGGCAGCAGGAGGGCGGCAGCTCTCGCAGGGCGCA
GGGCGGGCGGCCAGGATCATGTCCACCACCATGCCAAGTGGTGGCGTTCCCTCTGTCCATCTGGGGCT
GGCGCGCTGCATCGCGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGACTTCAGGCTTACCGAATGCAGGCC
TATTTCACCATCTGGGACTTCCAGGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCT
GGGTGCCATTGGCCTCTGGTATCCATCTTGCCTGAAATGCATCCGATTGGCAGCAGTGGAGGACTCTG
CCAAGCCAACATGACACTCACCCTCCGGATCATGTTCAATTGTCAGGTCTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACCTCTGGATGTCCACAGCTAACATGTACACCGGCATGGTGG
GATGGTGCAGACTGTTAGACCAGGTACACATTGGTGCAGGCCTGTGTTGGCTGGGCGTGGAGGCC
TCACACTAACATTGGGGTGTGATGATGTGCATGCCCTGCCGGGGCTGGCACAGAACAGAAACTACAAA
GCCGTTCTTATCATGCCCTAGGCCACAGTGTGCTTACAAGCCTGGAGGCTTCAGGCCAGCAGTGGCTT
TGGGTCACACCAACAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGCAGTGTGAATGCTCTAACACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAAACAAAGGAGATCCCCTAGATTCTTCTGACTCACAGCAGTGAAGTTAGAAAAGCCT
CGATTTCATCTTGGAGAGGCCAATGGCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTTCATCTCTTAAATTTAAACT
TTCTACTCTGATGAGAGAATCTGGTTTAATCTCTCTCACATTGATGATTAGACACACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCAGCTTATCCCAGAAAACCTTTGAGCTAGTAATAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGCTCTCCACCTTGAGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTGTAATCTCTCCAGGCCATGATCTGGTTTCTT
ACACTGTGATCTTAAAGTTACCAACCAAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTG
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCCCTGAGCTCTCCACTGGAGTCCTCTTCTG
CGGGGGTCAGAAATTGTCCTAGATGAATGAGAAAATTATTTTTAAAGTCCTAAATAGTTAA
AATAAAATAATGTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTGACATTGTCTATATGGTACTTGTAAAGTCATGCTTAAGTACAACATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCATTGGAGGCTGAGGAGGAAGGACTTGAAGGCCAGAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCCTACAAAATACAGAGAGAAAAATCAGCCAGTC
TGGTGGCATACACCTGTTAGTCCAGCATTCCGGAGGCTGAGGTGGAGGACTTGAAGGCCAGGGAGGT
TGGGGCTGAGTGAACACCACACTGCACTCCAGGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAATAATGAAACACAGCAAGTCCTAGGAAAGTAGGTTAAAACTAATTCTTAA

FIGURE 118

MSTTCQVVAFLLSI¹GLLAGCIAATGMDMW⁷QDLYDN¹⁰PVTSV¹³FQYEGLWRSCVRQSSGFTECRP
YFTI¹⁴LGLPAMLQAVRALM¹⁷IVGIVLGAIG²⁰LLVS²³I²⁴FALKCIRIGSMEDSAKANMTL²⁷TSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTL³⁰IGGVMMCIA
CRGLAPEETNYKAVSYHASGH³³SVAKP³⁵GGFKAST³⁷FGSNT³⁹KNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTTCTGCGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTCCGGAGTCC
AGCTGGCTAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGCTGTTCTTG
GTGGTGGAAATGGGGCACAGTGGCTGACTGTATGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATCCCTGCTGGCTTCTCCGGACCTACAGGCAGGCCAGAG
GACTGATGTGCTGCTCCGTATGTCCTCTGGCTTATGATGCCATCCTGGCATGAAATGC
ACCAAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTCAT
CATCACGGGCATGGGGCTCATCCCTGAGCTGGGTGCAATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCGACGGCA
CTGGTGCTGATTGTTGGAGGAGCTGTCCTGCTGGCTTTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACTTCCCACATCGCACAACCCAAAAAGTTACACACCGGAAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTATGTTTTAACTTTACTATAAAGCCATGCAAATG
ACAAAAAATCTATATTACTCTCAAATGGACCCAAAGAAACTTGATTTACTGTTCTTAACGCT
AATCTTAATTACAGGAACGTGCATCAGCTATTATGATTCTATAAGCTATTCAAGCAGAATGAGATA
TTAAACCCAAATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATTCTA
CTCTTTTATCTTACTTCAAATGACATTGCTAAAGACTGCATTATTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGAACATTATCTCACATAGAGACATGCTTATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAATAGAACTCAACTATTGCTTCAAGGGAA
ATCATGGATAGGGTGAAGAAGGTTACTATTAAATTGTTAAAAACAGCTTAGGGATTAATGCTCTCCA
TTTATAATGAAGATTAAATGAAGGCTTAATCAGCATTGTAACATTATCTCACATAGAGACATGCTTATGGT
GCTGTTTTAGCCTAGGAGTTAGAAATCTAACCTCTTATCCTCTCCAGAGGCTTTGCATTCAAACGTCT
CTTGCTGTTAAATTAAACATTAAAACGCAGATATTGTCAGGGCTTTGCAAGGGCTTGCTTCAAGCAGAATG
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGTGTTAGGAAAGTG
AAAATATTTTGTGTTGTTGAGTACAGACTTGAGGTTCATCAATATAAATAAGAGCAGAAAATA
ATATTTAATAAGTATTGAGTACAGACTTGAGGTTCATCAATATAAATAAGAGCAGAAAATA
TGTCTTGGTTTCTTGTGTTACCAAAAAACACAACAAAAAGTTGTCCTTGTGAAAATAATTCCTCTGTA
CCATTCTGTTAGTTTACTAAAATCTGAAATACTGTATTTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTCGACGTCTGTAGCTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAAATTGTACATTCTAATT

3. 4.5 5.6 6.7 7.8 8.9 9.10 11.12 13.14 15.16

FIGURE 120

MATHALEIAGLFLLGGVGMVGTVAVTVMQPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSSLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRGTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGC CGCGGGT GAAAGGCGCATTGATGCAGCCTGC CGCGGCCTCGGAGCGCGCGAG
CCAGACGCTGACCACGTT CCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCCTGCCAGCC
GGGAGCCATGCACCC CAGGGCCCGCCGCTCCCCG CAGCGGCTCCGC GGCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGAAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGT CAGACGGGAGCC TGGGCCAATGTTATTCCGGGTACACCTGGGATCC CAGGTGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCGAGGGAAAGCTTGAGGGAGTCCTGGACACCCAAC TAC
AAGCAGTGTTCATGGAGTT CATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTG TACATT
TACAAAGATGC GTTCAAATAGT GCTCTAAGAGTTTGTCAGTGGCTACTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTCACATTCAATGGAGCTGAATGTT CAGGACCTT CCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAAATTAAATTATCATCG
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGG
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTCTACTGGATGGAATT CAGTTCTCGC ATC
ATTATTGAAGAACTACCAAAATGGTTAATTTCATTGCTACCTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATT TAAATAAGTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAGTGTGATTCACACTGTTTAAATCTAGCATTATTCA TTTG
CTTCAATCAAAGTGGTTCAATATTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTGGAATATTGTTGTGGTCTTTGTTCTTAGTATAGCATT TTA
AAAAAAATATAAAAGCTACCAATCTTGTACAATTGTAAGAATT TTTTATATCTGT
TAAATAAAAATTATTCCAACA

FIGURE 122

MRPQGPAAASPQRLRGLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGTACGGGCTCTCGCCTCTGGCTTCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGAACTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCGCTGGAAGAATACATCATGTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTA
 ACCGGCCCCCTCCCCACCCCCAAAAAAACTGTAAGATGAAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTGATGTTGCTGCGAATGCGGTGTTGGGATTATTTGTTCTGGAG
 TGTTCTCGTGGCTGGCAAAGAATAATGTTCCAAATCGGTCCATCTCCAAGGGGTCCAATT
 TCTTCCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
 GCCCTAAGCCAAAGCAAAGACCTAAGGACGACCTGAAACAATACAAAGGATGGGTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGAGCACTGGTTATACCCCCACTGCTTACTGACAATG
 CTTCTCTGCGGAACGAGGATGCCCTAAGGGCTGAGGTGTGAAGGCAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCCCTAAGTATATCTGCTGGTTGCTTAGGTTGTCCCTCGCT
 ATAACAGCCTCAAAACTAAGTATAATCAATTAAAGGCTCAACCAGCTCACCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATTGACGAAATGTTTAATGGAATACGCAACTCAAAGA
 GCTGATTCTTAGTCCAATAGAATCTCTATTCTTAACAATACCTCAGACCTGTGACAATT
 TACGGAACTTGGATCTGTCTTATAATCAGCTGCAATTCTCTGGATCTGAACAGTTGGGGCTTG
 CGGAAGCTGCTGAGTTACATTACGGCTTAACCTCTGGAGAACCATCCCTGCGAATATTCCA
 AGACTGCGCAACCTGGAACCTTGGACCTGGATATAACCGGATCCGAAGTTAGCCAGGAATG
 TCTTGCTGGCATGATCAGACTCAAAGAACCTCACCTGGAGCACAATCAATTTCAGCTCAAC
 CTGGCCCTTTTCCAAGGTTGGTCAGCCTCAGAACCTTACTGCACTGGAAATAATCAGTGT
 CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTAACAAAGGCTGATTATCAGGCAATGAGA
 TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCGGAACTGCAAGCCCTCAACCTGGAT
 TCCAACAAGCTCACATTATTGGTAAGAGATTGGGATTCTGGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATGGGAATGCAGCAGAAATATTGCTCCCTGTAAACTGGCTGAAAGTT
 TTAAAGGTCTAAGGGAGAATACAATTATCTGTCAGCTCCAAAGAGCTGCAAGGAGTAATGTG
 ATCGATGCACTGAGAAACTACAGCAGCTGTCAGCTGGCAAAGCTCCCTGAGGAGGTTGATCTGGCAG
 GGCTCTCCAAAGCCGACGTTAACGCCCCAGGCGGAAGCATGAGAGCAAAACCCCTT
 TGCCCCCGACGGTGGGAGCCACAGAGCCGCCAGAGACCGATGCTGACCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGGCAGCGTGGCCTTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT
 CTACGTGTATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAAGCGCTCCCTCATGCAA
 GGCACAGGAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTATGTA
 GATTATAACCCACCAACACGGAGACCAGCGAGATGCTGTAATGGGACGGGACCCCTGCACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATGTAATGGCTTAAAGCTTAAAGCT
 GGGAAATAAGTGGGTGTTTATTGAACTCTGGTGAATCAAGGGAACGCGATGCCCCCCTCCCC
 TTCCCTCTCCCTCACTTTGGTGGCAAGATCCTTCTGTCGGTTTAGTGCATTATAACT
 GGTCAATTTCCTCTCATACATAATCAACCCATTGAATTTAAATACCAACATGTAAGCTT
 GAACTCCGGTTAATATAACCTATTGTATAAGACCCCTTACTGATTCCATTATGTCGCAATT
 GTTTAAGATAAAACTTCTTCATAGGTAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRТИPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKELHLEHNQFSKLNLA FPRLVSLQNLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLFIGQEILDWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKE LQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVL VILLVIYVSWKRYPASMKQLOQR
SLMRRHRKKKRQSLKQMTPSTQE FYVVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCCGTCGGAGGAGGAGGGCTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTCGCGGCTACCGTGGCCGAGCT
AGCAACCTTCCCTGGATCTCACAAAAACTCGACTCCAAATGCAAGGAGAACAGCTTTGCTC
GGTTGGGAGACGGTGAAGAGAAATCTGCCCCCTATAGGGAAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGGTCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAAC TGACCTAGTGAAGGTTAGATGCAAATGAAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCGTGGTGTACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGGCAGGCTGGGTACCCAAATACAAAGAGCAGCACTGGTGAATATGGGAGATTAA
ACCACTTATGATACAGTGAACACACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCATGTCA
TCAAAAGCAGAATAATGAATCAACCACCGAGATAAACAAAGGAAGGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTGAGTCTATATAAAGGCTTTTAC
ATCTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLRREVVFHKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMQUEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCGIVASILGTPADVIKSIRMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSLRLMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTGGCGCGCGCAGGAGAGCGGCCGGCGTCAGCTCTCGAC
CCCCGTGTCGGCTAGTCCAGCGAGGCGACGGGCGGTGGGCCATGGCCAGGCCGGCATGG
AGCGGTGGCGCACCGGCTGGCGTGGTGAACGGGGCCTGGGGGATCGGCCGCGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCACCCGGACTTGATCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTGGCCGGCTGACACCCCTGCTCTAGGCAGCACAGTGGTGGAA
GGACATGTTCAATGTAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACAGTCCATGA
AGGAGCGGAATGTGGACGATGGCACATCATTAACATCAATAGCATGTCGGCCACCGAGTGT
CCCCGTCTGTGACCCACTCTATAGTGCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCCAGGTGGTGG
AGACACAATTGCCCTCAAACCTCCACGACAAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTACTGTGGAGCTCC
TCCTTCCCTCCCCACCCCTTCATGGCTTGCCCTCTGCCCTGGATTTAGGTGTTGATTTCTGGAT
CACGGGATACCAACTTCCTGTCCACACCCGACCAGGGCTAGAAAATTGTTGAGATTTTATA
TCATCTGTCAAATTGCTTCAGTTGAAATGTGAAAAATGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGCCCCCTGGGACTTGGCCTTGTCTCAGTG
TCTCCCTTGACATGGAAAGGAGTTGGCCAAAATCCCCATCTTGTCTGACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCCTCACCTTATATCTGTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCTGCCCACTGCACCCCTCCCCCTATCTATCTCCTCTCGCTCCCC
AGCCCAGTCTGGCTCTGTCCCCCTGGGCTATCCCTCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTGATGGTACATGGTACATAAAAAAGAAAAATCGCAACCAA
AAAAAAA

FIGURE 128

MARPGMERWRDRLLALVTGASGGI GAAVARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLI
PYRCDSLNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHI ININSMMSGHRLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETOFAFKLHDKDPKEAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGGTGCCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCCTCCACCCCTCCTTGACTGCGGGCCGTTCAGGTGCAGAGTCTCAGTTGCCGGGAGC
ACCTCCCCCTCCGGAGGCAGGTCTGCTAGAGGGCCTGGCCAGATTCCAGTTCTGGTTTCATGC
CAGCCGTAAAGGCATGAACTTTGGGGAATCACCGATGCCATTAAAGAGGTTCTGCCA
GGATGGAAATGTTAGGTCGTCTGTGTCGCGTGGTCATTTCAGTAGCCACAGCCACTGTGG
CCGGTGAGTGCTGGAAATGAGGAACTGAGAAATTATTCTCATGTATTTCTCATTTATTTA
TTAATTTTTAACTGATAGTGTACATTTGGGGTACATGTGATTTGGATACATGTATACAA
TATTATGATGCAAAATCAGGGTACTGGGATATCCCACATCAAACATTTATTTTATTCTT
TTAGACAGAGTCTCACTCTGTACCCAGGCTGGAGTGCAGTGGGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCCCACCTCCAAGTAGCTGGACTACAGGCAT
GCACCACAATGCCCAACTATTTGTATTTTAGTAGAGACGGGGTTTGCCATGTTGCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACATCCACTTGCCCTGGCCCTCCAAAGTGTTGATTACA
GGCGTGAGCCCACCGGTGCCTGGCTAAACATTTATCTTTCTTGTGTTGGGAACTTTGAATTTAT
ACAAATGATTGTTACTGTATCTCCCTGCTGCTATGGAACACTGGGACTTCTCCCT
ATCTTACTGTATTTGTACCGATTAACCAACCGTACTTATCCCCATCCCTCTATCCTTCCC
AACCCTGTATCACCTCATTCTACTCTCTACCTCCCATGAGATCCACTTTTTTAGCTCCCACATGT
AGTAAGAAATGCAATTTGCTTCTGGCCTGGCTATTCACTTAACATGACTTCCTG
TTCCCATCCATGTGCTGCAAATGACAGGATTCGTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

MGLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRSVAREHLPSPRGSLLRGPRPRIPVLVSCQPV
KGHGTLGESPMFKRVFCQDGNVRSFCVCAVFSSHQPPVAVECLK

FIGURE 130

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCAACAG
GAATACCAGAGGGATGCTACAACTCTCACCTTCAGAACACAAATAAATAATGCTGGGATTCCCTCAGAT
TTGAAAAACTTGTGAAAGTAGAAAGAATATACCTATACCACAAACAGTTAGATGAATTTCCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGAGAATTACATTAGATGACAACCTCTGTCTGCACTGAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTCCCTGTCGGTAATCACCTTAGCACAATTCCCTGGGGTTT
GCCCAAGGACTATAGAAGAACTACGCTGGATGATAATCGCATATCCACTATTTCATCACCCTCTTCAAG
GTCTCACTAGTCTAAACGCCCTGGTTCTAGATGAAACCTGTTGAACAATCATGGTTAGGTGACAAAGTT
TTCTTCACCTAGTTAATTGACAGAGCTCCCTGGTGCGGAATTCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTTATCTCAAGATAACCACATCAATGGGTGCCCAATGCTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCATAAATAACCTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGTGATTCTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCGTGGGATGGCTATTAGGATCTCAATGCAGAACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACGTGCAATACCCAAACAGTGTACCTGCCCAGGACAGTGGCAGCTCAGTGAC
CAAACAGCCAGATATTAGAACCCCCAAGCTCACTAAGGATCAACAAACCACAGGGACTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACACCTGTGATACCATTCAATCTCTGGAAACTGCTTACCTATG
ACTGCTTGAGACTCAGCTGGCTAAACTGGCCATAGCCCGATTGGATCTATAACAGAAACAATTGT
AACAGGGGAAACGCACTGAGTACTTGGTCACAGCCCTGGAGCCTGATTGACCTATAAAACTATGCA
CCATGGAAACCAGCAACCTCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCT
CGAATGTACAACCCCTACAACCAACCCCTCAACTGAGAGCAAGAGAAAGAACCTTACAAAAACCC
AATTGCTGCCATATTGGTGGGCTGTGGCCCTGGTTACCATTGCCCTTGTGTTAGTGTGTTAGT
TTCATAGGAATGGATCGCTTCTCAAGGAACGTGCAATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGGAAATCAGGGAAACTCTTTCAAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTGTAATACACCCATTTCCTCTAAATGGAATGAATCTGT
ACAAAAACAACTACAGTGAAGGAGCTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGTGCTGAAGGACTCACAGCAGACTTGTGTTGGTTTTAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSPYL
EELHLDNSVSAVSIEEGAFRDSNYLRLFLSRNHLSLIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLDKVFFNLVNLTELSVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNRGRLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAPAQGQWPAPVTKQPD
IKNPKLTKDQTTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKGHPAFGSITET
IVTGERSEYLVTALEPDSPYKCMVPMETSNLYLFDETPVCIETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCAGAGTCCTTGCCAGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGCTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGCCATGGCCAGCCTGGGCTGCTGCTCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCCTGGCTGGACACTGCTGAAAGTAAAGCCACCATGAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAGTGTCCGGAGAAGTGGGCCAGGAGCCCTG
TGCAGCCGCTGAGCCTGCCGTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGTACCTAACGACTAACAGAGAGTCCAGCTGACCCCTCCAGCCCG
GTTTTGGAAGCTCCCACATGCCCTGGATCCACACTGATGCCCTTGGTGTACCCACGTTGCC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCCTGGTGCAGCTGCTGGGAACCGGAGC
GACAGCAGCAGGCCCTGCCCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCGTCCCACCAACTGCTCTTCCCTCTGGGCCAGAATGAGGGATGCAACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCAACCTACCCGGACATCTCATGGAAAACATCATGTTCTG
TGGAATGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTCAGGGAGCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGCAGAAAAACATTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTGCCTGGTTCAAGC
AATTCTCTGCCATCCTCCCGAGTAGCTGGACTACAGGAGCGTGCACCATACCTGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGCTCGAACCTCGT
CTCAAGAGATCCGCCAACCTCAGGCTCCAAAGTGTGGATTATAGGTGTGAGCCACCGTGTG
GCTAAAAGCATTCAAAGAGACTGTGTGAATAAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCCCTAGATGGCTGCTCTCCACAAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCCAGCAAACAGAGAGGCCACACCCATCCACACCG
CCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCCTGGAGGAGGAACAGACCC
TTTAGTCCTCATCCCTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCCGACACCCAACTTGGGAAGCCCTGAGTAGGCAGGGCAGGGTAGGTGGG
GGCCGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLARAQGAERVTSSEQRPAMASLGLLLLLTLPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGCTTCTTCCGTCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTCGGAGGTGCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTGGCGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTCCCTTAAG
ACAGATGGGAGTTTGTGGTCATGATATACTTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTGATCCCGTTCGAGTGGATATCAGTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGCAGACTGCCCTATCCTCTCCAAATGAAATCTCAGGT
CCACCTTCTTACTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTCTTCCCTTATTGATATTGTGCTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTCTGAGTTCATGACAAGACTCTCTTCAAAATCATGGCAAATCTAGCAGCGCAGCAG
TAAAACAGGCAAAGTGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGTGGCATCCAAGTCTTGGAAAACCAGCTTCATTTGTCTGTAT
GAGTCATCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTTGGTACACGAGAAAACCCAGCTTCATTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAATTAT
ATGAACTACTATACATTATGTATATTAATTAACATCTAATCCAGAAATCAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

Important features of the protein:

Signal sequence:
amino acids 1-23

Transmembrane domain:
amino acids 161-182

N-glycosylation site.
amino acids 184-187

Glycosaminoglycan attachment sites.
amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 151-154

N-myristylation sites.
amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.
amino acids 238-241

ATP/GTP-binding site motif A (P-loop).
amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTTCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCGGAGGGAGGCTATATCGTCAATTCCCCAAACAA
GTTTGACATTCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGAGACAATAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTCTTCACGGAGGCTTGGCAGT
TTTCTTACTCCTGTGGTCTCCAGATTCAAGGCTAAAGATGAAAGCTCTAGTCTTGCCTTCAGC
CTTCTCTCTGCTGCGTTTATCTCTATGGACTCCTTCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTTCAGGAAATCGAAATGGATTTCTGAGATAACGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAACTTAAAGGAGGACTGAGTCTTGCAGACACA
AAGCCTGCGAATCGATGCTGCCCTGCGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTACTCTCGGAAGATCAGCAGCCTGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGCATGCCAA
ACCACCATCTTTACTGTACTAGTCTGTGCTGGTCACAGTGTATCTTATTGCATTACTTG
CTTCCTTGATGATTGTCTTATGCATCCCAATCTAATTGAGACCATACTGTATAAGATTT
TGAAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTATTATTGCTATTAA
ATGTATTATTTTTACTGGACATGAAACTTTAAAAAAATTCAACAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTACAGTAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCATTGTATTCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGAT
ATTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTGATGTGAAATTGCAC
ATCTACCTTACAATTACTGACCATCCCAAGTAGACTCCCCAGTCCATAATTGTGTATCTTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTCTTGCATACCAAAAAAAA
AAAAAAA

FIGURE 138

MRFQPKTSFDISPMSFSIYSLQPVAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLDdrvFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAIIMTCHGEEAMKKYSQILSHFEKLEPQAAVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCGGAAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTTCCGCATCCAGCC
TAGCGTGTCCACGATCGGGCTGGGCTCAGGACTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTCCTCGGGGATCTTCCCGCTCCCGTCTGTTCCCTGCGAGACGGAA
CACGGAGCGGAGCCCCCAGCGCCGAACCTCGGCTGGAGCCAGTTCTAACCTGGACCACGCTGCC
ACCACCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCCTGAGAGATGATTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTTGAGGAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAACATCAAGGCATTGATGACGGGAG
CCTTCTGGTTTGTGACGCTCATCAGGAACCTCAATTCTCTGACTGCTGGAAAGACAGTGTGA
TAAGACAAGAAAAGCAGCTGGAAAAAGAATAGTCCTTATGGAGATGAAACCTGGGTTAAATTA
TTCCAAAAGCATTGTGAAATATGATGAAACACTCATTTCTGTCAGATTACACAGAGT
GGATAATAATGTCACGGGATTGGATAAGTATAAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGCTGGACCACATTGCCACATTCAAGGCCAACAGCCCCGATTGGGAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTACTGAGTCAGAGGAGAGA
GACGCCCTTACCAATTGCTGGTTTTGTGACCATGGCATGTCGAAACAGGAAGTCACG
GGGCTCCTCACCGAGGGTGAATACACCTCTGATTTAACCTAGTTCTCGCTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAGCAGCTCAATAGCGGATGTCGACACTGGGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGGCTCCTATTCCAGTTGAGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAGATCTGGTTTGAGCAGTTAAATGTCAGAAAGATGGCATGG
GAACCTGGATCAGACTGTACTTGGAGGAAAGCATTAGCAGTCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCACTGGATGCTCTGAGACGCTGAGCTGTCAGCTGGCTTGAGTCACAAGTGGCCCAG
TTCTCACCCCTGCTCTGCTCAGCGTCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAAGTCCCA
CTGTCATCTCCGGTTTCTCTCTTATTGGTGTATCTGGTTCTTCGGCCGTCACGT
CATTGTGCACTCAGCTGAAAGTCTGCTACTCTGTCGCTCTGTCGCTGGCTGGCGAGGCT
GCCTTCGTTTACAGACTCTGGTGAACACCTGGTGTGCAAGTGTGGCAAGTGCCTGGAC
AGGGGGCCTCAGGGAGGACGTGGAGCAGCCTTATCCAGGCCCTGGGTGTCGGACACAGGTG
TTCACATCTGTCGTGCAAGTCAGATGCCCTAGTTCTGGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGAGCTGCAGCTGAGGGAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTGCACCACTCATCTGCCACCCAGATGCATCCT
GCCTCATCAGGTCCAGATTCTTCCAAGGCGGACGTTCTGTTGAAATTCTTAGTCCTGGCC
TCGGACACCTTCATTGTTAGCTGGGAGTGGTGTGAGGCAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCC
ACCCCAACCTGCAAGGCCCTCATCCCTCTGGCTTGAGGCCCTGAGAGGCCCTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTCACTAGGGCACAGGCTCTCGGAGGCCAGGATGATCTG
CCACGCTGCACTCGGGCCATCTGGCTCATGCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTTACAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFPAPVRSSARAEHGAEPAPAEPSAGASSNWTLLPPPLF
SKVVIVLIDLALRDFVFGSKGVFKMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHVFVEYDGTTSFVSDYTEVDNNV
TRHLDKVLKRGDWDLILHLHGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGHDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGTGGCACTGATGTTACTGCTGCTGGAGTACA
ACTCCCTATAGAAAACA
ACTGCCAGCACCTTAAGACC
ACTCACACCTTCAGAGTGAAGAA
ACTTAAAC
CCGAAGAAAATTCA
GCATTCA
TGACCAGGATC
ACAAAGTACTGGC
CTGGACTCTGGGA
ATCTCAT
AGCAGTTCC
AGATAAAA
ACTACATAC
GCCAGAGA
GATCTTGC
ATTAGC
CTCATC
CCTGAGCT
CAGC
CT
CTGGAG
AAAGGA
AGTCC
GATTCT
CCTGGGG
CTCTAA
AGGGGAG
TTTG
CTCTAC
TGTGACA
AGGATA
AAGGACA
AAAGTC
ATCC
ATCC
TCAGCT
GAAGA
AGGAG
AAACT
GATGAAG
CT
GGCTGCC
AAAGGA
ATCAGC
ACGCC
GCC
CTTC
ATCTT
TATAGGG
CTCAGG
TGGG
CTCTGG
A
ACATG
CTGGAG
TCGG
GGCTC
ACCC
GGATGG
TTCAT
CTGC
ACCT
CCTG
CAATTG
TAATG
AGC
CT
GTTGGGG
TGAC
AGATA
AAATT
TGAGA
ACAGG
AAACAC
ATTG
AATT
TCATT
CAACC
AGT
TTG
CAA
AGCTG
AAATG
AGCCCC
AGTG
AGGT
CAGCG
ATTAG
GAAACT
GCC
CATTG
AACGC
CTTC
CTCG
CTA
ATTG
GA
ACTA
ATTG
TATA
AAA
AC
CAA
AC
CTG
CTCA
CT

FIGURE 142

MLLLLEYNFPPIENNQCHLKTHTFRVKNLNPKKFSIHQDHKVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMS PSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-9

Interleukin-1
amino acids 37-182

FIGURE 14.3

CTAGAGAGTATAAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAA
TCCAGGATCCTGCTCTTCTGTCCTGTAGGAGTGCCTGTTGCCAGTGTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGCTCAATTAACTCCTGTGGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGCTGGCGCTGCTGCTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTCCACAGAGCAAAGGACTGTGATGAACACAATGTCATGGCTCGTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCTGGAAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAATTGAAGACGACATTGACAACGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTACCTGCTTCTCACCATCAGCACCAAGGCCCTGGATGACTCAGTCAGC
CTCCTGAACAAGACCTGCTGGAGGGATTCCACTTGAGTGAACCCACTCACAGGCTGTCCATGT
GCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGATCATAAACATCAGCATTTAAGAAAAAAAAAAAAAAA
AAA

MLGLPWKGGLSWALLLGSQILLIYAWHFHEQRDCDEHNMARYLPATVEFAVHTFNQQSKDY

FIGURE 144

MLGLPWKGGLSWALLLGSQILLIYAWHFHEQRDCDEHNMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGCTGGCCACTGTCTGCATGCTGCTCTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAACATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTTCATCAAGCAAGGCCAAGCTCGACATT
GACTTCGGAGCCAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAAGGGGAGTTCCAGAACGCCAGAACACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTGGTGGAGAGGGCGCAGGACT
TCGGGTCAACATGCACCAAGCCAGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCAGTGCAGGCCACCTGCTCACAGTACTTCCAAACAACCTTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCAGC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCCTGGCTTGCTAACCCAGGTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGTACTTATGCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK

LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQLWLWRLVQELCSILKHCEFWLERGAGLRTMHQPVLLCLLALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACTGTCTGGTCCAGAGTCTCATTT
CCTGATGATTTAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTCTCTGGCCTCCTCT
CTGTCTTCTTCCCTTTCTTCTTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCACTGGGGTAGGTCACTGAGTCTTAGTTTATTTTGAAATT
CAACTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFILLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEEDKIKKTYPPENKPG
QSNYSFVDNLNLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIKEKVTPMAAIQDGLAKGENDETGSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKKLEKNATDNISKLFPAPSEKSHEETDSTKEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCGTGGAGCTCAA
GATGGCTCTGAGTGGGGCGTGTGCTTCCGAATGAAGGACTCGCATTGAAGGTGCTTATCTGC
 ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAAGGTATTAAGGGTAAGAGATCAGC
 GTGGTCCCAATCGGTGGCTGGATGCCAGGCTGTCCCCGTATCCTGGTGTCCAGGGTGAAG
 CCAGTGCCTGCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
 AGCTCTATCTTGGTGCAAGGAATCCAAGAGCTTCACCTTCTACCGGGGGACATGGGGCTCACC
 TCCAGCTTCGAGTGGCTGCCAACCCGGCTGGTCTGTGCACGGTGCCTGAAGCCGATCAGCC
 TGTCAGACTCACCCAGCTTCCCAGAATGGTGGCTGAATGCCCATCACAGACTTCTACTTCC
 AGCAGTGTGACTAGGGCAACGTGGCCCCAGAACCTCTGGCAGAGCCAGCTGGGTGAGGGT
 GAGTGGAGGAGACCCATGGCGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
 TGGCACCTGACCACTTTGCTTCTGGTCCAGTTGGATAAATTCTGAGATTGGAGCTCAGT
 CCACGGTCTCCCCACTGGATGGTCTACTGCTGTGGAACCTGTAAAAACCATGTGGGTAAA
 CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGAGTGGTGGGAATCTCCTGCT
 TAATGGTAACTGACAAGTGTACCCCTGAGCCCGCAGGCCAACCCATCCCCAGTTGAGCCTTATA
 GGGTCAGTAGCTCTCCACATGAAGTCTGTCACTCACCCTGAGGAGAGGGAGGTGGTCATA
 GAGTCAGGGATCTATGGCCCTGGCCAGCCCCACCCCTTCCCTTAATCTGCCACTGTCATA
 TGCTACCTTCTCATCTCTCCCTCATCATCTTGTGTGGCATGAGGGAGTGGTGTAGTCAGAA
 GAAATGGCTGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAACCAA
 GATACAATCAAATCCCAGATGCTGGTCTTATTCCATGAAAAGTGTCTGACATATTGAGA
 AGACCTACTTACAAAGTGGCATATATTGCAATTATTAAATTAAAAGATAACCTATTATATT
 TCTTATAGAAAAAGTCTGAAAGAGTTACTTCAATTGTAGCAATGTCAGGGTGGCAGTAT
 AGGTGATTTCTTTAATTCTGTTAATTATCTGTATTCCTAATTTCTACAATGAAGATGA
 ATTCCCTGTATAAAATAAGAAAAGAATTATCTGAGGTAAGCAGAGCAGACATCATCTGA
 TTGCTCTCAGCTCCACTTCCCAGAGTAAATTCAATTGAAATCAGGCTCTGCTCTGGTGG
 TTGAGTAGTGTGAGTCAAGGAAACAGATCTCAGCAAAGGCACTGAGGAGGAGCTGTGAGTTGT
 GTGGCTGAAATCTGGTAAAGGAACCTAAAGAACAAAATCATCTGGTAATTCTTCTAGAAG
 GATCACAGCCCCGGGATTCAAGGCATTGGATCAGTCTCTAAGAAGGCTGTGACTGGTGA
 ATTGTGCCCCCTCAAATTACATCCTCTTGAATCTCAGTCTGTGAGTTATTGGAGATAAG
 GTCTCTGAGATGTAGTTAGTTAAGACAAGGTCTGGATGAAGGTAGACCTAAATTCAATAT
 GACTGGTTCTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
 AAGATGAAGGCAGAGATGGAGTTTGAGGCCACAAGCTAAGAACACCAAGGATTGTGGCAACC
 ATCAGAACGTTGAAGGCAAAGAACATTCTCAGACTTCCAGCTCTGAACGAAGAAAGAATAAATT
 CTGGCTGTTAACTCTCAGACTTCCAGCTCTGAACGAAGAAAGAATAAATTCTGGCTGTTAA
 GCCACCAAGGATAATTGGTACAGCAGCTTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
 CTCTCGTGTGTTACATTCTGTGTGTCCTCCACAATGTACCAAGTTGCTGTTGTGACCAA
 TAGAATATGGCAGAAGTGTGAGGCATGCCACTTCAAGATTAGGTATAAAAGACACTGCA
 GCTTCACCTCTCTGCCACCCACCGCCCCAATCTATCTGGCTCACTCGCTCTGGGG
 AAGCTAGCTGCATGCTATGAGCAGGCTATAAAGAGACTTACGTGGAAAAATGAAGTCTCCT
 GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGGATAATCGATGTTGTGTTT
 AAGTTGCTCAGTTGGTCTAATTGTTATGCAGCAATAGATAAATAATGCAAGAAAGAG

FIGURE 152

MVLSGALCFRMKDALSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGG
QCLSCGVGQEPHTLIEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFCLTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTTTCCCTATGGGGACCCCTGGCCACCAGCTGCCCTCTCTTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATACCCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTGGCTGATAAA
CAACACAGACGTTCGTCTCATGGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAGCAACAGGCTAACGACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTAAACCAAAAGGAAGATGGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTCGTGTAGTTACAAAGGAAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTCTAACGATAGATATTGATAACATTCACTGTAACGGTGTTC
TATACACAGAAAACAATTATTTTAAATAATTGCTTTCCATAAAAAGATTACTTCCAT
TCCTTGTAGGGAAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATAAGACTGCATTATTTATCATTATTAAATATGGATTATTAT
AGAAACATCATCGATATTGCTACTTGAGTGTAAGGCTAATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKSVSSFIMGTLATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRILIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQUEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTAAAATAAAATCAGGACTCTAACCTGCTCCAGTCAGCCTGCTCCACGAGGCCTGT
CAGTCAGTGCCGACTTGTGACTGAGTGTGCAGTGCAGCATGTACCAAGGTAGTCAGAGGGC
TGCTGAGGGCTGTGCTGAGAGGGAGAGGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTGGGCTGGGGCCAAGTGGAGTGAGAAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCATTAGCCTTCTACAGGTGGTGCAT
TCTTGGCAATGGTATGGAAACCCACACCTACAGCCACTGGCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCAGTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTAACAGCAGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGCTCCCCCAGGACCTGTACCACGCCGT
TGCCCTGTGCCGCACTGCGTCAGCCTACAGACAGGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTTGTGTGTGCAGGCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCTGAAGGGCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTGTGGAGCAG
CAGGATCCGGGACAGGATGGGGGCTTGGGAAACCTGCACCTCTGCACATTGAAAGGTTTCAA
CAGCTGCTGCTAGGGCCGGGAAGCTGGTGCCTGTCAATTCTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACCTCTGTCTTCCCTTTCCATCCCCGCTACCCCTG
GCCAGCACAGGCACATTCTAGATATTCCCCCTGCTGGAGAAGAAAGGCCCTGGTTTATT
TGTGACATGGATGATTCTGAGGAGGAAGCTTATTGAATGTATAGAGATTATCCAAATAAATAT
CTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMV р MGTHTYSHWPSCCP SKQDTSEELL RWSTV PVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**ATG**TCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGACTATTCA
ATTTGATGAATGTAAGCTGGTACTCCGGGAGATGCCAGCATCCGTTGAAGGCCACCAA
GATTGTGTGACGGGCAAAGCAACTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTCTGGTGTAAATGGACATTTCTACATCGGCTCCCTGTA
GAGCTGAACACAGTCTATTCTATTGGGGCCATAATATTCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCGTGAATTTCACCTCACAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACAACCAACTCCCTGGAAACAGATACATGGCTTATCCAACACACCAC
TATCATCGGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAACACGCGAGCTTCAGTGGTGA
TTCCAGTGACTIONGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTCTACTGTGGC
AGCGACTGCATCCGACATAAGGAACAGTTGTGCTCTGCCACAAACAGGCCTCCCTCCCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCCTCCTCTGCTGTCCTGCTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTCCCTT
TCTACCACCAACTACTGCCCTCATTAAGGTTCTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTCACTGAATTCTCAAAACATTGAGAAGTGGATCATTGAA
AGTGGCAGAAAAGAAAATAGCAGAGATGGTCCAGTGCAGTGGCTGCCACTAAAAGAGGCA
GCAGACAAAGTCGTCTCCCTCTTCAATGACGTCACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGGAGAACTCTCAAGACCTCTCCCTGCCCTTAACCTTTCTGCA
GTGATCTAAGAAGCCAGATTCACTGCACAAATACGTGGTGGTCACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTGCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTCTCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAGCCTGCCACG
ATGGCTGCTGCCCTTGT**AG**

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNVS梧
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEGDGPSMSVNFTSPGCLDHIMKYKKCVKAGSLWDPNITACKNEETVEVNFTTTPLCNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTVGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVFPPLDNNK
SKPGGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSIECFHHTICYFTEFL
QNHCRCSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristylation site.

amino acids 116-122

Amidation site.

amino acids 488-492

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCCTGCATGGCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
GAGACCCCTGGTGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLTVGCTCVTPVIIHHVQ

Signal sequence:
amino acids 1-30

N-glycosylation site.
amino acids 83-87

N-myristoylation sites.
amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGGATTC
CAGGCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCTTGGGGGGGCCAGCAC
AGGGCCTCAGGCTGGTGCACCTGAGACAGAAGATGCCTGTGCCCTGGTTCTGCTGTCC
TGGCACTGGGCCAGAAGCCCACTGGTCTTCTCTGGAGAGGCTGTGGGGGCCAGGACCTAAC
CACTGCTCCGGGCCCTCCTGCCGCTCTGGGACAGTGCACACTCTGCCCTGCCCTGGGGACAT
CGTGGCTGCTCCGGGGCCCGTGTGGGCCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTGCGTGTGGTGTCCACTTGGCGTGCATGGGACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGGGTGGAGGAGCCTAGGAA
TGCCTCTCCAGGCCAAGTCGTGCTCTCCCTCCAGGCCAACCTACTGCCCGTGCCTGCC
TGGAGGTGCAAGTGCCTGCTGCCCTGTGCAAGTTGGTCACTGTGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTATACTCAGGCCAGGTACGAGAA
GGAACACTAACACACAGCAGCTGCCCTGCCCTGGCTCAACCTGTCAGCAGATGGTGCAC
ACGTGCATCTGGTTCTGAATGTCCTGAGGAGCAGCACCTCGGCCCTCCCTGTACTGGAAATCAG
GTCCAGGGCCCCCAAAACCCCGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGA
CCACACAGACCTGGTTCCTGCCCTGTATTGCACTGGACCGCAGATCATTACCTTGA
CGAACATCTGCCCTTCAGGGAGGACCCCGCGCACACCAAGAACCTGTGGCAAGCGCCCGACTG
CGACTGCTGACCCCTGCAAGAGCTGGCTGTGGACGCACCGTGTGCTGCCCGCAGAACGGCACT
GTGCTGGGGCTCCGGTGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCTGGAGAACG
TCACTGTGACAAGGTTCTGAGATTCCATTGCTGAAAGGCCACCTAACCTGTGTTCAAGGTG
AACAGCTGGAGAAGCTGCACTGCAGGAGTGCTGTGGCTGACTCCCTGGGCTCTCAAAGA
CGATGTGCTACTGTGGAGACAGGAGGCCAGAACAGATCCCTGTGCTTGGAACCC
GTGGCTGACTTCACCTACCCAGCAAGCTCCACAGGGCAGCTGCCCTGGAGAGTACTTA
CAAGACCTGCACTGCCAGGGCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAAAGCGCTGGGCCCTCGTGGCTGCCCTACTCTTGGCG
CTGCGCTTCTCCCTCATCCTCTCAAAAAGGATCACCGCAGGGTGGCTGAGGCTTGTGAAA
CAGGACGTCCGCTGGGGCGCCAGGGCGCAGGGCGCAGGGCTGTGCTCCCTACTCAGCCGATGA
CTCGGGTTTCGAGGCCCTGGTGGCGCCCTGGCTGCCCTGTGCCAGCTGCCGCTGCGCTGG
CCGTAGACCTGTGGAGGCCGCTGTGAACTGAGCGCGCAGGGGCCCTGGCTTGGTTACGCCAG
CGGCCAGACCCCTGCAAGGAGGGCGCGTGGTGTCTTGTCTCTCCGGTGCCTGGGCT
GTGCAGCAGTGGCTACAGGATGGGTGTCCGGGCCGGCGCACGGCCCGCACGACGCCCTCC
GCCGCTCGCTCAGCTGGCTGTGCCGACTTCTGCAAGGGCGGCCGGCGACTACGTGGG
GCCCTGCTCGACAGGGCTGCTCCACCCGGAGCCGCTACCCGCCCTTCCGACCGCTGCCGCTT
CACACTGGCCCTCCAACTGCGCAGACTTCTGGGGGCCCTCGAGCAGGCCCTGCCGCTTCCG
GGCGCTCCAAAGAGAGGCGAGCACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CATCCCCGGGACTCCCGCGCCGGACGGGGGGTGGGACCAGGGGGGGACCTGGGGGGGGGG
CGGGACTAAATAAGGAGCAGCCTGTTTCTAAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVGPQDATHCSPGLSCLRWDSDILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCRLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLFSQAYPTARCVLLEV
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNV
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQ
LWQAAARLRLTLQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKLEFPPLLGHPNLCV
VNSSEKLQLQECILWADSLGPLKDDVLLLETRGPDNRSLCALEPSGCTSLSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLLLLKKDHAKGWILRLKQDVRS
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRQTLQEGGVVVLLFSP
GAVALCSEWLQDGVSQPGGAHGPDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFR
LPSQLPDLGALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVVPGAGPAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCACCTTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAATTCCAGTCCAGCAACTTTGA
AAACATCTGACGTGGGACAGCGGGCAGAGGGCACCCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGTGGCAAAGAAGGGCTGTCAAGCGGATCACCCGGAAGTCCCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTCAACCGCT
GTCAGTGCAGGGAGGCGGTCAGCCACCAAGATGACTGACAGGTTCAAGCTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATGATTCAAGATGATTGTTCATCCTA
CCCCCACGCCAATCCGTGCAAGGGCATGCCAACCGGCTAACCTGGAAAGACATCTTCCATGACCTG
TTCTTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCCACCTGGAGGGAAAGCAGAGAGA
ATATGAGTTCTCGGCCTGACCCCTGACACAGAGTCTTGGCACCATCATGATTGCGTTCCA
CCTGGGCAAGGAGAGTGCCTTACATGTGCCAGTGAAAGACACTGCCAGACCGGACATGGACC
TACTCCTCTCCGGAGCCTCCGTGTTCTCCATGGGCTCCTCGCAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGCCGCTGCACCTCCAACTCCCTGAACTGCCAGCGAGTCTGACTT
TCCAGCCGCTGCCTCATCCAGGAGCACGTCCTGATCCCTGTTGACCTCAGCGGGCCAGC
AGTCTGGCCAGCCTGTCAGTACTCCAGATCAGGTGTCCTGGACCCAGGGAGCCGAGGAGC
TCCACAGCGGCACTGGCTGTCAGTACCTACTTAGGGCAGCCAGACATCTCCATCTCCAGC
CCTCCAACGTGCCACCTCCCAAGTCCCTCCCTAGTCCCTATGCCCAAAACGTCGCCCCGAG
GTGGGGCCCCCATCTGCACCTCAGGTGACCTGGCTATGGAGGAATCCAAGAAGCAAATCATGACCAGC
GCCATCTCTAAGGTCCAGGCTTCCCTCTATGCCCTCAAGCCACTCGGACAGCTGGCTCCCT
CCTATGGGTATGCATGGAAGGTTCTGCCAAAGACTCCCCACTGGACACTTTCTAGTCCTAA
CACCTTAGGCCTAAAGGTCAAGCTCAGAAAGGCCACAGCTGGAAGCTGCATGTTAGTGGCCT
TTCTCTGCAGGAGGTGACCTCCCTGGCTATGGAGGAATCCAAGAAGCAAATCATGACCAGC
CCCTGGGATTGACAGACAGAACATGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCAGCTCCCTCCTCTCAGTCCAGATCGAGGGCCACCCATGTC
CCTCCCTTGCAACCTCCTCCGGCATGTTCCCTCGGACCAAGGTCCAAGTCCCTGGGCC
TGCTGGAGTCCCTGTGTCAGGATGAAGGCCAAGGCCAGGGCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACTGGATTCTCTTTCAGAGGCTGGCCCTGACTGTGCAAGTGGGAGTC
AGGGAAATGGGAAAGGCTGGTGTCTCTCCCTGTCCTACCCAGTGTACATCCTGGCTGTCA
ATCCCCTGCCTGCCATGCCACACACTCTGCACTGGCTCAGACGGGTGCCCTGAGAGAAC
AGAGGGAGTGGCATGCAGGGCCCTGCATGGGTGCGCTCTCACCGAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTCTGGGAGCAGCTGTGAGACAAGCGCTGCTCGCTGAGCCCTG
CAAGGAGAAAATGACAGTGCAGGAGGAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGTCTGAGGAAATTTGCTCTCCCTGCCCCATTCTGGCAGTTTC
ACAATCTGCTGCAGAGACATGAGGCCCTGCCTCTGCTATTGTCAGGTTCAAAGGTGGAAAGAGA
GCCTGGAAAAGAACCAAGGCCAGGGAAAGAACAGAAGGGCTGGGCCAGAACAGAACACTGC
ACTTCTGCCAAGGCCAGGGCCAGGGACGGACTCTAGGGAGGGTGTGGCTGAGCTCA
TTCCCAAGGCCAGGGCAACTGCCGTGACGTGCACTGGCTCAGATTCTGCTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCTTTCTGCAGGCAGGGTTCAGACCCCT
ATCTGAGAATGGGTTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
TGTAATGATGTCACAACCTTGCAAGCTGCTGCCCTGGTTCAAGCCATCTGGCTCAAATTCCAGC
CTCACCACTCACAAGCTGTGACTTCAAACAAATGAAATCAGTGCCTGAGAACCTCGGTTCTC
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTAAAGTCTTAATAGTGCCTGGTACATGGGAGTGCCTGAGAACCTCGGTTCTC
AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTEYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNQQRVLTFQPLRFIQEHLVLPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAIISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESLVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCGCGTGGCCACAACATGG
CTGGCGGCCGGGCTGCTTCTGGCTGCTGGCTGGGGCGCTGGTGGTCCCAGCAG
TCGGATCTCAGCCACGGACGGCGTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGAAAGCTTGAAGACTTCAGGCCCTGATTGCGTTGTGAATTAA
AAAAAGGTGACGATGTATATGCTACTACAAACTGGCAGGGGATCCCTGAACCTGGCTGGA
AGTGTGAACACAGTTGGATATTTCACAAAGATTGATCAAGGTACTTCATAAAACACGGAA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTGTCTGCTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGAAGACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAAATCTCCTGAGGAGTCTGGGGGCGTGA
ACTTGACCCGTGCGCTGAGCCGAGGCATTAGCAGAGCTGATTAGCAGAGGATGGAGAAGGTGCTTCT
CAGAGAGCACCAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGGCCACCCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCAGGGAAAGCAGAACAGACTGGCAATAGTTCTCCTGCCTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATCAAATCTAAGTTGTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTGGCAGTTGCATTTGAGCCAAACAAAAATATTATTTCCCTCTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFLVGLALWWVPGQSDLSHGRFRSDLKVCGDEECSMILMYRGKALEDFTGPDCRFVN
FKKGDDVVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYEEELHIPADEDFVCFEGGRD
DFNSYNVEELLGSLELEDSPVEESKKAAEVSQHREKSPPEESRGRELDPVPEPEAFRADSEDEGEGA
FSESTEGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVCIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGGAAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTTCTGCTTCGCCCTAGGACATACACGGGACCCCCTAACCTTC
AGTCCCCAAACCGCACCCTGAAGTCTTGAACCTCAGCCCCCACATCCACGCGCGCACAGG
CGCGGCAGGCGCAGGTCCCGCGAAGGCGATGCGCGCAGGGGTCGGCAGCTGGCTCGGGC
GGCGGGAGTAGGGCCCGCAGGGAGGAGGCTGCATATTCAAGAGTCGCGGGCTGCCCTG
GGCAGAGGCCGCCCTGCCACGCAACACCTGCTGCTGCCACCGCGCAGCTGGCTGG
TCTCGCTGCTGCTGGCGCGCGCTGCTGCGGCCACGGAGCCTCTGCCGCCGCTGGTCA
GCCAAAAGGTGTGTTTGTGACTTCAGCAAGCTCCCTGCTACAAAATGCCCTACTTCCATGA
GTCAGCCGAGTGAGCTTCAGGAGGCAGCCTGGCTGTGAGAGTGAGGGAGGAGTCCTC
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAA
ACAGGGATTCGTGATGGTGTGATTCTGGATAGGGTTGGAGGAATGGAGATGGCAA
ACAGGGATTCGTGATGGTGTGATTCTGGATAGGGTTGGAGGAATGGAGATGGCAA
ACAGGGATTCGTGATGGTGTGATTCTGGATAGGGTTGGAGGAATGGAGATGGCAA
TGCCTGCCAGATCTTACCAAGTGGCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTTCTCGGAAGTGAAGAAGTGTGTTGTGATGTATACCAACCAACTGCCAATCTGGC
CTTGGGGTCCCTACCTTACCAAGTGGAAATGATGACAGGTGTAACATGAAGCACAATTATTTG
CAAGTATGAACACAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTATCTACAAATCAAC
GAGACACCCATCAGAATGTTGTTACTGAAGCAGGTATAATTCCAATCTAATTGTT
ATACCAACAATACCCCTGCTTACTGATACTGGTTGCTTGGAAACCTGTTCCAGATGCT
GCATAAAAGTAAGGAAGAACAAAATAGTCAAACCCAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAACTCATTGACTTGGTTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACATAGCTGGAATGGCTGAAATCAAAGGATCTG
AAAGATGAACTGTAAGCTCCCCCTGAGGCAAATATAAGTAATTGTTATATGCTATT
TTAAAGAATATGCTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAAGATAAAGTGTGTTATCAACAC
GTGGAGTAATGCTGTTAGAAGCAATTCTTTATTCTCCTACCTTCTACAGTGTG
TGTATATTGTTAGAATTACAGTGTGCAAAAGTATTCTACCTTGTGATAAGTGT
ATGAACCTGTTCTAATATTATTGTCATCTCATTTCAATACATGCTCTTGTGATTAAAG
AAACTTATTACTGTTGCAACTGAATTCAACACACACAAATATAGTACCA
AGAAAAGTTGT
TTCTCGAAATAATTCTACCTTCAGCTCTGCTTGGTCAATGCTAGGAAATCTCTCAGA
AATAAGAAGCTATTCTACAGTGTGATAAAACCTCCTAAACATTTACTTAGAGG
GAGGAT
TGCTAATTCTACAGTGTGCAAGACATGTGCTTATAATTATTGTTAGCTTAA
AAACTAAAGTAACTTTGTTAATAGGTGCAAAACACTAATGCA
TGACATACACAAATAATCATATGCTTCACACGTTGCTATATAATGAGAAGCAG
CTCTG
GGGTCTGAAATCAATGTTGCTCTTCTGCCCCTAAACAAAGATGGTTGTT
ATTGACACTGGAGGCAGATAGTTGCAAGTTAGTCTAAGGTTCCCTAGCTG
ACTATATTAGTATAACAAAGAGGTATGTGTTGAGGACAGGTGA
ACAAGCACAGCACACAGACATTAGGAAGGAAAGGA
ACTACGA
AACTTAAAGAGTCTTCTCAATGCTTACGCCCTGTTCTCTCAAGAGAAAGTTG
AACCCATCAGTGATGCGATATTGATGAGGGTTGCTT
CTGCTTATCTCTAGTTCTCAATGCTTACGCCCTGTTCTCTCAAGAGAAAGTTG
CTGCTTCTCATATGTCCTGCTCTTAAACCAATAAAGAGTTCTGTTCTGGGGAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLGAALLCGHGAFCRRVSGQKVCFADFKHPCYKMAFHELSSRVSFQEARNLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIVVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217